

X-RAY CRYSTALLOGRAPHIC STUDIES

of some NATURAL PRODUCTS

being a

T H E S I S

presented for the Degree of

D O C T O R of P H I L O S O P H Y

in the

FACULTY of SCIENCE

of the

U N I V E R S I T Y of G L A S G O W

by

STRUTHER ARNOTT, *B.Sc.*

Chemistry Department  
University of Glasgow.

December 1959.

ProQuest Number: 13850700

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 13850700

Published by ProQuest LLC (2019). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

The author wishes to express his appreciation to the following persons and institutions for their assistance and cooperation during the course of the work reported in this thesis:

**ACKNOWLEDGEMENTS.**

*The Author wishes to declare his indebtedness to Professor J. Monteath Robertson, F.R.S., his first inspiring mentor in the study of X-ray crystallography and the supervisor of the researches reported in this work; to Alexander Davie, Dr. George Sim and David Watson, who shared part of the labour; and to the Gourock Ropework Company Limited who provided a scholarship.*

---

## PREFACE.

This thesis describes researches carried out by the author (some of them in collaboration with others) under the supervision of Professor J. Monteath Robertson, F.R.S., in the X-ray Crystallographic Section of the Chemistry Department of Glasgow University from October 1956 to December, 1959.

It is concerned with the determination of crystal structures using the techniques of X-ray crystallography. In it the variety of these techniques is outlined by their enumeration in the first chapter. The second chapter describes the successful application of crystallographic methods to the solution of the structure of limonin, the bitter principle of citrus fruits. Limonin has been known for some 127 years, and has resisted the efforts of many chemical schools throughout the world to establish its constitution and configuration by the methods of classical organic chemistry. It has been described by Professor D.H.R. Barton as the last remaining major structural problem in classical natural product chemistry.

The third chapter indicates the efforts to use the same investigating tools to increase the structural knowledge of a new class of antibiotics called macrolides. These efforts have not yet led to a definitive conclusion and thus Chapter (III) must wear the same unsatisfactory character that most incomplete crystallographic researches possess.

A number of tables summarising results are interpolated in the text which is supplemented by four Appendices. The first of these deals with the computing aids used in the work described; the second is a list of calculated structural amplitudes and phase-angles for the epi-limonol iodoacetate structure at its last refinement stage; the third contains all the figures mentioned in the text; and the fourth shows structural formulae of major interest to the work undertaken.

The work on limonin was largely a team effort. The early work thereon was undertaken by the author and has been published (Arnott & Robertson, 1959). The two dimensional work was shared with D. Watson and G.A. Sim and for the three-dimensional work the team was joined by A.W. Davie. All aspects of the work were intimately shared except for the computations of the a-axis projections and Harker section which were done by D.W. while the b-axis projections and attempts to correlate these with FI was largely the author's responsibility.

A preliminary report of the limonin structure is to appear in *Experientia* in February, 1960.

---

TABLE of CONTENTS

	<i>Page</i>
<i>Acknowledgments.</i>	
<i>Preface.</i>	
<i>Chapter I. The determination of the chemical structure of organic molecules by X-ray crystallographic methods.</i>	
1.1. <i>Introduction</i>	1
1.2. <i>X-ray scattering and crystal structure</i>	2
1.3. <i>The solution of the phase problem</i>	4
1.3.1. <i>Direct Methods</i>	4
1.3.2. <i>Chemical Crystallographic Methods</i>	5
1.3.3a. <i>Trial and Error</i>	6
1.3.3b. <i>Fourier Transforms</i>	6
1.3.3c. <i>Heavy atom method</i>	6
1.3.3d. <i>Isomorphous replacement method</i>	9
1.3.3e. <i>Use of anomalous scattering</i>	9
1.3.3f. <i>Patterson methods</i>	10
1.4. <i>The refinement of crystal structures</i>	12
1.4.1. <i>The <math>F_o</math> synthesis</i>	13
1.4.2. <i>The method of Least Squares</i>	13
1.5. <i>References</i>	15
<i>Chapter II. The structure of limonin</i>	
2.1. <i>Preamble</i>	16
2.2. <i>The structural problem</i>	16
2.3. <i>The preliminary crystallographic survey of limonin and its derivatives</i>	18
2.3.1. <i>Introduction</i>	18
2.3.2. <i>Limonin and its solvates</i>	19
2.3.3. <i>Salts of limonin</i>	23
2.3.4. <i>Salts of limonilic acid</i>	23
2.3.5. <i>Haloacetic esters of epi-limonol</i>	24
2.4. <i>Collection and processing of the experimental data</i>	
2.4.1. <i>Collection of intensity data for crystals of epi-limonol iodoacetate</i>	27
2.4.2. <i>Estimation and correction of intensities</i>	30

	<i>Page</i>
2.4.3. <i>Determination of film-to-film scaling constants</i>	33
2.5. <i>The determination of the structure of limonin</i>	35
2.5.1. <i>The 2-D investigation</i>	36
2.5.2. <i>The Harker section at <math>y = \frac{1}{2}</math></i>	37
2.5.3. <i>The 3-D investigation</i>	39
2.5.4. <i>The refinement of the structure</i>	43
2.6. <i>The nature of the limonin structure</i>	53
2.7. <i>Conclusion</i>	55
2.8. <i>References</i>	56
 Chapter III. <i>The preliminary investigation of the structure of the macrolide antibiotics</i>	
3.1. <i>Introduction</i>	58
3.2. <i>The X-ray crystallographic investigation of the     macrolide antibiotics</i>	59
3.2.1. <i>Macrolides and macrolide derivatives</i>	59
3.2.2. <i>Crystal data for pykromycin and neomethymycin         methylene dichloride solvate</i>	60
3.2.3. <i>Neomethymycin methylene dichloride solvate</i>	60
3.3. <i>References</i>	64
 Appendix I. <i>Computation of Fourier Series</i>	65
II. <i>Structure amplitudes and phase-angles for     epi-limonol iodoacetate</i>	72
III. <i>Diagrams and Illustrations</i>	80
IV. <i>Structural Formulae</i>	93

## CHAPTER I.

The determination of the chemical structure of organic molecules by X-ray crystallographic methods.1.1. Introduction

Although it is not the purpose of this work to make a minutely critical assessment of the methods of chemical structure analysis, it has been considered desirable to outline the various approaches available in the X-ray crystallographic method of structural study which was used in the work later described.

Structural analysis in the classical manner is based on deductions made from a series of chemical reactions involving the unknown molecular species and its derivatives. In carbon chemistry in particular, experiments have to be devised to determine not only the molecular constitution but also to establish the configuration of the atoms at any asymmetric centre which may be present. The information from purely chemical reactions has, in recent years, been increasingly supplemented by the results of spectroscopic examination in the infra-red and ultra-violet regions and other physical methods. Such means of investigation all depend on relating the chemical and physical properties of the unknown substance to those of previously determined structures. They are essentially indirect compared

with the methods of studying the sub-microscopic nature of matter by considering the interactions between matter and electromagnetic radiation or particle beams. Just as the scattering of alpha-particles by metal foils led to the discovery of the general nuclear structure of atoms, so the diffraction of X-radiation by crystals has led to the solution of problems of molecular geometry. It is this second phenomenon which has been utilised in this work.

## 1.2. X-ray scattering and crystal structure.

To use the scattering of X-rays by crystals in structure determination, it is necessary to relate X-ray diffraction patterns to the atomic arrangements in the crystals. While the conditions for the production of diffracted beams are determined solely by the dimensions of the crystal lattice, the intensities of the beams are controlled by the atomic arrangement.

The structure factor,  $F(hkl)$ , of a crystal plane whose Miller indices are  $h$ ,  $k$ , and  $l$  is the result of combining the waves scattered by the whole contents of the unit cell and has the form:

$$F(hkl) = \sum_j f_j \exp 2\pi i(hx_j + ky_j + lz_j) \quad \dots\dots 1.1$$

where  $(x_j, y_j, z_j)$  are atomic coordinates and  $f_j$  is the scattering factor for the  $j$ th atom. This scattering factor can be resolved into two parts:

$$f_j = f_j^0 \exp(-B \sin^2 \theta / \lambda^2) \quad \dots\dots 1.2$$



where  $f_j^0$  is the atomic scattering factor which is a function of the scattering angle,  $\theta$ , and depends on the distribution of electrons in the atomic species in question. It can be calculated when the radial electron distribution function for an atom is known. Atomic scattering factors so derived by several workers have been collected in the International Tables (1935). The modifying exponential function in equation 1.2 is the temperature factor which takes account of the effect of the thermal motions of atoms in the crystal.

Equation 1.1 implies that all the scattering matter in a crystal unit cell is concentrated in a number of spherically symmetrical atoms situated at points  $(x_j, y_j, z_j)$ . It requires that for each atom the electron distribution function is known so that the atomic scattering factor can be calculated, and further that some estimate of the temperature factor should be available. Therefore it is sometimes more convenient to use a more general expression for the structure factor:

$$F(hkl) = (V/abc) \int_0^a \int_0^b \int_0^c \rho(x, y, z) \exp 2\pi i (hx + ky + lz) dx dy dz \dots \quad 1.3$$

where  $\rho(x, y, z)$  is the density (in electronic units) of the scattering matter at any point  $(x, y, z)$  in the unit cell,  $V$  is the unit cell volume,  $a$ ,  $b$ , and  $c$  the lengths of the unit cell edges.  $|F(hkl)|$  is simply the ratio of the amplitude of the radiation scattered in the order  $(h, k, l)$  by the contents of one unit cell to that scattered by a single electron under the same conditions.

Since a crystal can be considered as a triply periodic and continuous distribution of electron density it can

be represented by a triple Fourier series:

$$\rho(x, y, z) = \sum_{-∞}^{+∞} \sum_{-∞}^{+∞} \sum_{-∞}^{+∞} A(pqr) \exp 2 \pi i (hp+kq+lr) \dots 1.4$$

It can be shown by combining equations

1.3 and 1.4 that

$$\rho(x, y, z) = \sqrt{V} \sum \sum \sum F(hkl) \exp 2 \pi i (hx+ky+lz) \dots 1.5$$

$$= \sqrt{V} \sum \sum \sum F(hkl) / \cos 2\pi (hx+ky+lz) - \alpha(hkl) \dots 1.6$$

where  $\alpha(hkl)$  is the phase angle given by

$$\alpha(hkl) = \tan^{-1} (B/A) \dots 1.7$$

where A and B are respectively the real and imaginary parts of the complex number  $F(hkl)$ .

It is in the solution of equations 1.1 or 1.3 that the possibility of structure determination lies. The difficulty arises from the fact that only the moduli of the Fourier waves can be found from the X-ray diffracted intensities which contain no direct information about the phases of the waves.

### 1.3. Solution of the phase problem

This section seeks to outline the principal methods by which the phase problem can be overcome and the determination of a crystal structure achieved from intensity data and a minimum of other information. The heavy atom and Patterson techniques which were of greatest use in the later analysis of limonin will be discussed in greater detail.

1.3.1. Direct Methods. While direct solution of equations 1.1 is possible only when both the real and imaginary components of the  $F(hkl)$ 's are known, it is generally possible

to solve the equations

$$|F(hkl)|^2 = \sum_{n,m}^N f_n f_m \exp 2\pi i \sum h (x_n - x_m) \quad \text{for}$$

$(x_n - x_m)$  etc., and from these solutions to determine the atomic co-ordinates. Efforts begun by Ott (1927) and Avrami (1938) to derive general procedures for the solution of such equations have not been attended by success sufficient to influence the practice of crystal structure analysis.

A second direct method depends on the recognition that the positivity of electron density and similar items of direct space information are sufficient to impose severe limitations on the moduli and phases of structure factors. Such limitations are of most use in the case of centro-symmetric electron density distributions where the value of  $\alpha(hkl)$  cannot be any value between zero and  $2\pi$ , but has either of the discrete values  $0 (+)$  or  $\pi (-)$ , which is equivalent to signing the structure amplitudes in the manner indicated. The mathematical expression of these limitations in the form of inequality relations between structure factors is particularly associated with the names of Harker and Kasper (1948) while Sayre (1952) and Goedkoop (1950) have derived equality relations.

1.3.2. Chemical crystallographic methods. These usually involve procedures by which good approximations to the phase angles are derived and then combined with the observed structure amplitudes to enable an approximation to the true electron density distribution to be calculated (using equation 1.6). From this, new atomic co-ordinates are obtained which are used to calculate better phase-angles and the whole

procedure reiterated until a complete and accurate structure is obtained.

1.3.2a. The method of 'trial and error' requires a fairly extensive prior knowledge of the chemical structure. This knowledge is utilised in placing atoms and molecules in plausible positions in the unit cell until sufficiently good agreement between calculated and observed structure amplitudes allows the methods of Fourier analyses or least squares refinement to lead to the final accurate representation of the structure. This method, requiring so much previous structural information, is more suited to investigations whose aim is the derivation of accurate interatomic distances and angles than to researches whose primary interest is the determination of a largely unknown structure.

1.3.2b. The method of Fourier transforms requires a similar intimate preknowledge of molecular structure. The Fourier transform of an atom or group is a function representing the amplitude and phase of radiation scattered in a particular direction by the group. It can be evaluated from a knowledge of relative atomic positions in a molecule and used to determine the position of the atomic group in the crystal unit cell by being orientated with respect to the crystal reciprocal lattice until a fit with the observed structure amplitudes is obtained.

1.3.2c. It has been by using the 'heavy atom' technique (Robertson, 1935, 1936; Robertson and Woodward, 1937, 1940) that most of the triumphs of X-ray crystallographic structure determination have been achieved. In this method initial

approximations to the phase-angles are provided by a small number of atoms of dominant scattering power whose positions in the unit cell can be determined by Patterson methods.

The assumption is that the phase-angles derived in this way can be combined with the observed structure amplitudes in a Fourier synthesis yielding an approximation to the electron density distribution from which the sites of further atoms can be obtained. Better approximations to the phase-angles can then be calculated and the process continued until the whole structure is elucidated.

The structure factor for a crystal with one heavy atom in the unit cell can be written

$$F(hkl) = \sum f_H \exp 2\pi i (hx_H + ky_H + lz_H) + \sum f_j \exp 2\pi i (hx_j + ky_j + lz_j) \dots 1.8$$

If the heavy atom scattering factor,  $f_H$ , is much greater than any light atom scattering factor,  $f_j$ , then the first term will be greater than the second. It is not necessary that  $f_H > \sum f_j$  since the second term is a summation involving a number of atoms whose contributions may be opposed to one another. It is of interest to consider what relative scattering power is required by the heavy atom for success in this method to be assured. The mean value of the ideal intensity is given by

$$I(hkl) = |F(hkl)|^2 = \sum_n f_n^2 \dots 1.9 \text{ (Wilson, 1949)}$$

If an atomic scattering factor,  $f_n$ , can be considered to be proportional to the atomic number of the species in question, and if it is considered desirable that the scattering power of the heavy atom should be at least equal to the collective

scattering power of the lighter atoms, than a lower limiting case would exist when  $Z_H^2 / \sum Z_j^2 = 1$  ( $Z_j$  is atomic number of  $j$ th atom and  $Z_H$  that of the dominant scatterer). The fact that in general the relative scattering power of a heavy atom is greater than that indicated by its atomic number suggests that  $Z_H^2 / \sum Z_j^2$  can have values rather less than unity without impairing the success of the method. In most of the published successful structure determinations the ratio has had values of greater than unity, but in the noteworthy case of the successfully solved vitamin B12 structure problem (dominant scatterer used was -Se CN) the ratio had a value of 0.47 indicating that unity was certainly not the useful lower limit (Crowfoot Hodgkin et al., 1955).

This method is attractive because of the possibility it offers for a structural determination to be effected requiring no more chemical information than the molecular formula. It is capable of great versatility in its application since it is but rarely impossible to prepare a suitable heavy atom derivative of a structurally unknown or little-known molecule. Particular accidents can destroy this generality of applicability: In the case where only one heavy atom is present in the asymmetric crystal unit of a non-centred molecular arrangement such as exists in space group  $P2_1$  the phase-angles calculated on the basis of the heavy atom will be those which obtain in a centrosymmetric case, and any approximate electron density function calculated from them will show two superposed arrangements one of which is spurious. When this occurs (as in the case

of cholesteryl iodide - Carlisle & Crowfoot, 1945) considerable difficulty in resolving the two images is usually experienced. Similar additional pseudo-symmetry may be introduced when the heavy scatterer(s) lie(s) in a special position in the unit cell, or when, by the chance of it, independent heavy atoms have co-ordinates simulating additional symmetry.

1.3.2d. In the isomorphous replacement method the phase of the structure factor is determined by considering the diffraction patterns of two isomorphous heavy atom derivatives. It has the advantage over the simple heavy atom method in that comparison of the intensities from the two structures results in a more accurate initial phase determinations which at least shortens the refinement process. In addition higher accuracy can be achieved in the final refinement, since it can be performed for the lighter molecule of the isomorphous pair where absorption errors are smaller and the phase-angles more sensitive to the contributions of the light atoms. Unfortunately, in the non-centred case there exists an ambiguity in the determination of  $\alpha(hkl)$  which may require the structure to be disentangled from the image of its spurious inversion. In effect the method is similar to the heavy atom technique since one of the replaceable atoms requires to be a more dominant scatterer than the other atoms so that its position can be determined. However problems can be solved where the electron content of the replaceable atom is considerably less than that normally required in the case of the heavy atom method.

1.3.2e. Another method of solving the phase problem

for non-centric crystals has been described (Pepinsky, 1956a, b and 1957) which utilises the anomalous scattering of X-radiation by an atom whose absorption edge is near the value of the wavelength of the radiation employed (see James, 1950). When this anomalous dispersion occurs the atomic scattering factor becomes complex,  $f_j = f_j + i f_j'$ , and the normal Friedel's law equality,  $|F(hkl)| = |F(\bar{h}\bar{k}\bar{l})|$ , no longer holds. Pepinsky (loc. cit.) shows how the differences,  $\Delta = |F(hkl)|^2 - |F(\bar{h}\bar{k}\bar{l})|^2$ , can be used to solve the phase problem. The obvious drawback is that the anomalous scattering must be sufficiently pronounced for, the  $\Delta$ 's to be measured with reasonable accuracy. The same phenomenon was used by Bijvoet (1942) in a rather different way. He first solved a crystal structure by standard methods ignoring the fact that  $|F(hkl)| \neq |F(\bar{h}\bar{k}\bar{l})|$  after which pairs of  $F(hkl)$  and  $F(\bar{h}\bar{k}\bar{l})$  were examined to permit the assignment of absolute configuration to the molecule in question.

1.3.2f. The Patterson function of an electron density distribution  $\rho(x, y, z)$  is defined as:

$$P(u, v, w) = \int \int \int \rho(x, y, z) \rho(x+u, y+v, z+w) V dx dy dz \dots \quad 1.10$$

It can be represented by the Fourier series

$$P(u, v, w) = V^{-1} \sum_h \sum_k \sum_l |F(hkl)|^2 \exp(-2\pi i(hu+kv+lw)) \dots \quad 1.11$$

and has the enormous advantage of being independent of phase.

Patterson (1935) has shown that the summation corresponds to the total of all the interatomic vectors in the crystal structure. These are extremely numerous for a structure of any complexity and it may be impossible to disentangle them



except when prominent maxima representing vectors between heavy atoms are present.

Harker (1936) first pointed out that the properties of axes or planes of symmetry in a crystal concentrates important information in particular sections, the so-called Harker sections, of the Patterson function. The commonest symmetry element available for this purpose is the two-fold screw-axis ( $2_1$ ). If two atoms  $(x_1, y_1, z_1)$  and  $(x_2, y_2, z_2)$  occupy sites in the space group  $P2_1$  ( $\underline{b}$  unique axis), among the vectors from these atoms will be ones accounting for maxima in the Patterson function at  $(2x_1, \frac{1}{2}, 2z_1)$  and  $(2x_2, \frac{1}{2}, 2z_2)$  so that evaluation of the Patterson function at  $y = \frac{1}{2}$  could be expected to show peaks at  $(2x_1, 2z_1)$  and  $(2x_2, 2z_2)$ . There can be accidental complications for if,  $y_1 = y_2$  there will be peaks of type  $(x_1 + x_2), (z_1 + z_2)$  on the level  $y = \frac{1}{2}$ .

A further extension of the Patterson method has been suggested by Robertson & Beevers (1951) and by Buerger (1951). Since the Patterson function can be considered as a superposition of the images of the crystal structure in each of its component atoms, a plot of the Patterson function with its origin changed to a known atomic site in the unit cell will have maxima at other (unknown) atomic sites. If a number of such plots can be made with several heavy atoms in the unit cell as origins, and from these plots a function derived whose value at any lattice point is the minimum

value at the corresponding lattice points in the plots, then the maxima in this function will have a high probability of representing the remaining atomic sites in the cell.

#### 1.4. The refinement of crystal structures.

The refinement of a crystal structure is the process by which the best values of the atomic positions in the unit cell can be determined after fairly accurate co-ordinates for the atomic sites have been established. Refinement procedures are particularly important when the aim of an investigation is the determination of precise interatomic bond lengths and angles.

It has been found convenient to express the overall agreement between a postulated structure and the experimental data in terms of a reliability index,  $R = \sum | |F_o| - |F_c| | / \sum |F_o|$ , where  $|F_o|$  and  $|F_c|$  are respectively the observed and calculated structure amplitudes. While  $R$  is a useful guide in deciding whether any change in parameters has produced improved agreement with the observed data, its value in isolation is no measure of a structure's acceptability. Plausible bond lengths and angles and the association of the correct electron density with each atomic site are more fundamental criteria for assessing the validity and accuracy of a structure determination.

The following sub-sections describe briefly the two standard methods employed in the refinement of the structure of epi-limonol iodoacetate once initial atomic positions had been established by the early Fourier and Patterson syntheses.

1.4.1. The  $F_0$  Synthesis. In this method the electron density function is evaluated at suitable intervals (say 1/60 of unit cell edges) using the observed structure amplitudes and the phase-angles calculated from the approximately determined structure. Points of maximum electron density are then found by graphical interpolation or by using the method of Booth (1948) to calculate the peak position from values of the electron density at points near the summit. New phase angles can be calculated using the newly determined co-ordinates and the whole process reiterated until no further decline in  $R$  is observed.

There are disadvantages inherent in this method. In the first place a Fourier series is rarely complete in practice, but is generally terminated by the experimental conditions while its coefficients are still appreciable. Booth (1946) has suggested that this may be partly corrected by employing (in an  $F_C$  synthesis) the calculated structure factors in a Fourier synthesis in which terms omitted in the original  $F_0$  synthesis are left out. The co-ordinates derived from this new calculation will show deviations from those previously obtained; these deviations, with signs reversed, are the errors in the original co-ordinates. A further objection arises in  $F_0$  synthesis because it is difficult to decrease the influence of the less accurate coefficients on the result.

1.4.2. The method of least squares. This last objection can best be overcome by the method of least squares

(Hughes, 1941; Booth, 1947). If the errors in the measured structure amplitudes follow the normal or Gaussian law, the best agreement is obtained by minimising the function

$$R^1 = \sum w(hkl) (|F_o(hkl)| - |F_c(hkl)|)^2$$

where  $w(hkl)$  is a weighting factor inversely proportional to the probable error in  $F_o(hkl)$  and the summation is over all independent terms which lie within the limiting sphere. The atomic parameters on which  $R^1$  depends include quantities like the temperature factor as well as the atomic parameters. Each structure factor supplies one observational equation from which suitable linear equations are developed and reduced to the appropriate number of normal equations which can be solved by ordinary methods. Reiteration of the procedure leads to parameters which minimise  $R^1$ . Strict adherence to the requirement that  $w(hkl)$  should depend on an objective estimate of the standard deviation of the observation is not essential, since any reasonable weighting system will lead to co-ordinates which differ by amounts small compared with the random errors in the final co-ordinates.

Such a procedure will usually yield the best agreement between the structure formulated and the observed data but will not necessarily give a better account of the true structure. This is because the structure proposed will generally be conceived in terms of spherical atoms with each atom of the same species having equal scattering power. These are circumstances which rarely obtain.

1.5. REFERENCES

- AVRAMI, M., (1938), *Phys. Rev.*, 54, 300.
- BOOTH, A.D., (1946), *Proc. Roy. Soc. A*, 188, 77.  
 (1947), *ibid*, 190, 482.  
 (1948), 'Fourier Techniques,' p. 64 (C.U.P.)
- BLJVOET, J.M., (1949), *Koninkl. Neder. Akad.*, 52, 313.
- BUERGER, M.J., (1951), *Acta Cryst.*, 4, 531.
- CARLISLE, C.H., and CROWFOOT, D., (1945), *Proc. Roy. Soc. A*, 184, 6.
- CROWFOOT HODGKIN, D. et al., (1955), *Nature*, 176, 325.
- GOEDKOOP, J.A., (1950), *Acta Cryst.*, 3, 374.
- HARKER, D. (1936), *J. chem. Phys.*, 4, 831.
- HARKER, D., and KASPER, J.S., (1948), *Acta Cryst.*, 1, 70.
- JAMES, R.W. (1950), 'The Crystalline State', Vol II, (Bell, Loudon).
- OTT, H. (1927), *Zeit. Krist*, 66, 136.
- PEPINSKY, R., and OKAYA, Y., (1956a), *Proc. Nat. Acad. Sci.*, 42, 5.  
 (1956b), *Phys. Rev.*, 103, 1645.  
 (1957), *ibid*, 108, 1231.
- ROBERTSON, J.H., and BEEVERS, C.A., (1951), *Acta. Cryst.*, 4, 270.
- ROBERTSON, J.M., (1935), *J.C.S.*, 615.  
 (1936), *ibid*, 1195.  
 and WOODWARD, I. (1937), *ibid*, 219.  
 (1940), *ibid*, 36.
- SAYRE, D.M., (1952), *Acta. Cryst.*, 5, 60.

## CHAPTER II.

The Structure of Limonin.2.1. Preamble.

It is intended to describe in this chapter the evolution and successful conclusion of the determination of the structure of limonin (~~it~~) through an X-ray crystallographic study of the iodo-acetate ester<sup>(1)</sup> of epi-limonol: the nature of the structural problem will be outlined and the basic chemical facts known at the outset of the project enumerated; the sequence of operations leading to the discovery of the structure and the progress of its refinement will be described; some speculations on the biogenetic significance of the limonin structure solved by X-ray methods will be included.

2.2. The Structural Problem.

Limonin is a neutral, alicyclic bitter principle found in the fruits of citrus species and several other Rutaceae. It was first isolated by Bernay (1841) and so named. A number of subsequently isolated natural products - evodin from *evodia rutaecarpa* (Katmatsu, 1902), dictamnolactone from *dictamnus albus* (Thoms, 1923), obakulactone from *phellodendron amurense* (Fujita, 1931) - were shown in 1940 (Schichter and Haller), to be identical with limonin and it is from this date that intensive investigation of the structure by classical chemical methods commenced.

By the beginning of the present investigation

limonin had been characterised as having a molecular weight of 470, molecular formula  $C_{26}H_{30}O_8$ , melting point  $303-5^{\circ}C$ , and possessing optical activity (in acetone  $[\alpha]_D^{20.5} = -129^{\circ}$ .) Most investigators had assigned four of the eight oxygen functions to two lactone rings whose corresponding free carboxyl groups were determined titrimetrically to have acid strengths ( $pK_a^s$ ) of 2.7 and 4.7 in water (Emerson, 1952). The fifth oxygen function was sited in a  $\beta$ -substituted furan ring (Fujita & Hirose, 1954; Kubota & Tokoroyama, 1957; Corey et al., 1958) a feature which served also to account for the unsaturation in limonin. The sixth oxygen function had been characterised as ketonic by the formation of an oxime and a 2:4-dinitrophenylhydrazone (Kefford & Chandler, 1951). Examination of the infra-red spectrum of limonin showed that this carbon oxygen is attached to a six (or higher) membered ring. This carbonyl group could be reduced to either of two epimeric limonols; reduction by the Meerwein - Ponndorff method yields limonol (Chandler & Kefford, 1951,) while borohydride or sodium amalgam reduction gives epi-limonol (Melera et al., 1957; Fujita & Hirose, 1954, 1956). The last-named method of reduction is known to produce the more stable of two configurations (Barton, 1953) and accordingly limonol and epi-limonol were assigned axial and equatorial configurations respectively, a situation later confirmed by the X-ray analysis of epi-limonol iodoacetate. Neither the Zerewitinoff method of detecting active hydrogen atoms (Koller & Czerny, 1936) nor the analysis of the infra-red absorption spectrum of

limonin ( Rosenfeld & Hoffmann, 1951), indicated the presence of hydroxyl groups and accordingly the two remaining oxygen functions were considered to exist in two cyclic ethers (Geissmann & Tulegin, 1946).

The mode of formation of tetrahydrolimonin (in which saturation of the furan ring has occurred) and hexahydrolimonic acid (where hydrogenolysis of a lactone ring has also occurred) led to the conclusion that the allyl oxygen of the lactone was secured allylicly to the furan ring (II). The additional oxygen shown (II) plausibly explained the abnormal acid strength of hexahydrolimonic acid (Melera et. al., 1957).

It could be inferred from the molecular formula that the remainder of the molecular skeleton had to be bicarbocyclic. A Kuhn-Roth estimation of the number of methyl groups present suggested they were three in number, while a zinc dust distillation of limonin followed by selenium dehydrogenation had yielded 1:2:5 - trimethyl naphthalene (Koller & Czerny, 1936).

### 2.3. The Preliminary Crystallographic Survey of Limonin and its Derivatives.

2.3.1. Introduction. This section describes the search for limonin derivatives suitable for use in a structure determination by the methods of X-ray analysis and contains a report of the X-ray crystallographic data obtained for some of them (See Table 2.1). The unit cell parameters



for the crystals of each substance named in Table 2.1 were obtained by the measurement of zero-layer Buerger precession camera photographs taken with Mo  $K_{\alpha}$  X-radiation ( $\lambda = 0.7107\text{\AA}$ ). Thin films of zirconium metal were used to monochromatise the X-ray beam. The back of the camera film-holder contained two pinholes of known separation by which two pairs of fiducial spots could be printed on each film in a manner which gave correcting factors for each set of layer lines to allow for film shrinkage during processing. Derivation of lattice parameters from the film data also involved an instrumental proportionality constant  $F$  which was confirmed to be accurately 60 cms. by measurements using a calcite crystal as standard.

2.3.2. Limonin and its solvates. Limonin crystallises from methanol, ethanol, acetone, methylene dichloride, methylene dibromide and bromochloromethane in large six-sided plates in the orthorhombic system when the hot concentrated solutions are allowed to cool, or when the cold solutions are allowed to concentrate by evaporation. Within the general description provided the crystal habit varies from solvent to solvent, increasing molecular weight and involatility of solvent is accompanied by considerable thickening of the plates, so that in the case of methylene dibromide solutions the crystals recovered are tablets, hexagonal in plan and with bevelled edges.

Three solvates of limonin have been reported:  
limonin acetic acid solvate crystallises with a prismatic habit

from acetic acid solutions of limonin (Geissmann & Tulegin, 1946); limonin crystallises from solutions in methanol - methylene chloride in two metastable solvate forms, the more stable of which Emerson (1948) characterised as a methylene dichloride solvate, and the less stable as a hemi - (methylene dichloride) solvate.

Incomplete data for the optical properties of limonin has been published by Winchell (1943) and this has been elaborated by the work of Jones and Palmer (1949) who examined not only limonin but also its acetic acid solvate and the more stable of Emerson's methylene dichloride solvates. They determined that all three substances crystallised in the orthorhombic system class 6 (etch figures confirmed the enantiomorphism). They also determined the unit cell parameters and space groups for the first two crystal species (their results are those shown in parenthesis in Table 2.1). Their meagre conclusions on the structure of the limonin molecule were that it was comparatively flat (the c-axis was short) and possessed no extensive conjugation (refractive indices were low).

The unit cell dimensions and space groups of limonin and its acetic acid solvate were redetermined (see Table 2.1) and results obtained in good agreement with those previously reported. In both cases the only diffraction conditions were that (hoo), (oko) and (ool) were present only for  $h = 2n$ ,  $k = 2n$ , and  $l = 2n$  respectively (space group  $P2_1 2_1 2_1 - D_2^4$ ). Crystal densities were obtained by

measuring pycnometrically the densities of aqueous potassium iodide solutions in which the crystals could be suspended without either sinking or floating to the surface. These densities indicated in both cases a unit cell content of four formula weights.

The complex nature of a crystallographic structure determination for limonin involving as it would the siting of at least thirty-four crystallographically independent carbon and oxygen atoms in the unit cell made it obvious that procedures such as the 'heavy atom' or 'isomorphous replacement' methods would have to be utilised if useful image-seeking functions were to be calculated. Accordingly interest devolved on the structure of Emerson's methylene dichloride solvate whose content of two heavy scatterers (the two chlorine atoms) was deemed to have sufficient scattering power  $2Z_{Cl} = 0.38 \sum (27Z_C^2 + 8Z_O^2)$  to provide - once their positions in the unit cell had been determined - sufficiently good approximations to the correct phase-angles to allow Fourier and other image-seeking methods to be undertaken with fair chance of success. Since the reported instability of Emerson's solvate would be a considerable drawback, at the same time as Emerson's experiments were repeated a search for analogous and, it was hoped, more stable complexes involving halogenated methanes was undertaken.

Consideration of the X-ray diffraction spectra of crystals grown from solutions of limonin in methanol or in any of the pure halo-methanes available showed that they were

invariably unsolvated limonin. However when limonin was allowed to crystallise from methanol - methylene dichloride solutions, two crystal species could be obtained: on the surface of the solution there developed the characteristic hexagonal-shaped plates of pure limonin, while in the body of the solution there appeared prismatic lathes of square cross-section and invariably of a length at least ten times their thickness. These latter crystals rapidly became opaque even in contact with their mother-liquor, and when dry and exposed to the air for a few days became polycrystalline pseudomorphs. From no other methanol-halomethane solvent combination was any crystal species other than pure limonin recovered. The lathe-shaped crystals were assumed to be the more stable of Emerson's two solvates and the observation of Jones and Palmer that they belonged to the orthorhombic system was confirmed by the determination of their unit cell parameters and space group (see Table 2.1). The density corresponded to a unit cell content of four limonin and four methanol molecules. That the included molecular species was methanol and not methylene dichloride was confirmed by carbon, hydrogen and halogen analysis.

The hope of basing a structure analysis on Emerson's solvate was thus thwarted, but since it was thought desirable to show that a complex chemical structure might be solved by X-ray diffraction methods using no more chemical information than the molecular formula, it was decided to seek halo-acetic acid complexes analogous with the stable limonin acetic acid

solvate before considering 'heavy atom' derivatives involving functional groups in the unknown structure. No success attended efforts to crystallise limonin solvates from either chloro- or bromo-acetic acid melts, and although any solvent combination in which acetic acid was present yielded the acetic acid solvate, no corresponding compounds were obtained from methanol-haloacetic acid solutions of limonin.

2.3.3. The salts of limonin. There now followed efforts to prepare suitable derivatives using the chemical functions known to be present in limonin. The titration experiments of Emerson (1952) suggested the preparation of alkali metal salts of the potential carboxyl groups of limonin.

Accordingly, methanolic solutions of limonin were treated with slightly excess the theoretical volume of 6N sodium and potassium hydroxide. From the clear reaction mixtures the salts of limonin could be precipitated as granules by the addition of excess acetone. It was not found possible to grow single crystals of these salts whose concentrated solutions were invariably very viscous and yielded glass-like solids when the solvent was removed. Optical examination of a sample of one of these glasses showed that it was still isotropic after two years exposure in the laboratory.

2.3.4. Salts of limonic acid. Aqueous ethanolic solutions of limonic acid (III) were refluxed

for 24 hours with excess rubidium carbonate and then allowed to concentrate by evaporation. Syrups were obtained from which only crystals of hydrated rubidium carbonate were recovered.

2.3.5. Halo-acetic esters of epi-limonol. The ketonic carbonyl function of limonin can be reduced by borohydride in a stereospecific way yielding epi-limonol from which a chloroacetic ester can be prepared by the action of chloroacetyl chloride. The corresponding iodoacetate can be prepared by treating the chloroacetate with potassium iodide solution but no equivalent reaction occurs with bromide ions nor can the bromoacetate be prepared directly because of incidental bromination at other sites in limonin. Samples of the two halo-acetic esters were provided by Professor D.H.R. Barton and both were recrystallised from aqueous acetone solutions. The crystal habit of the former is long prismatic needles, and of the latter thin flaky plates with in the (001) face well developed. The determination of their unit cell parameters showed that both belonged to the monoclinic system. For both the only diffraction conditions were that (oko) was present only for  $k = 2n$ , hence their space groups were either  $P2_1 - C_2^2$  or  $P2_1/m - C_{2h}^2$ . The latter was rejected as being incompatible with the optical activity of both substances. The chloroacetate was known to be a mono-hydrate and this was confirmed by density measurements which required a unit cell content of two ester and two water molecules. The

iodoacetate possessed four molecules in the unit cell i.e. two in the asymmetric crystal unit.

Choice of the iodoacetate for further intensive structural investigation was decided by the following arguments: In linear absorption of Mo  $K_{\alpha}$  X-radiation the chloroacetate had an advantage by a factor of six, but this was more than offset by the greater phase-determining power of iodine over chlorine  $-Z_{Cl}^2 = (0.17) \sum z_L^2$  ;  $Z_I^2 = (1.74) \sum z_L^2$ . (It was unfortunate that the bromoacetate was not available for, since  $Z_{Br}^2 = (0.74) \sum z_L^2$ , a satisfactory compromise between absorption and scattering requirements would have been achieved; moreover since the absorption edge for bromine is near the wavelength of Cu  $K_{\alpha}$  X-radiation it might have been possible to use the phase-determining methods available when anomalous dispersion occurs). In addition the greater unit cell content of the iodoacetate (two molecules in the asymmetric crystal unit compared with one in the case of the chloro-derivative) would require the determination of double the number of independent atomic positions. This disadvantage is more apparent than real: the two chlorine atoms in the unit cell of epi-limonol chloroacetate are necessarily related by the screw axis of space group  $P2_1$ . In this space group the origin may lie anywhere on the screw axis and there would be no loss of generality if it were chosen to be midway between the two chlorine atoms. Approximate phase-angles calculated from the chlorine atom co-ordinates would have values 0 or  $\pi$ ,

Table 2.1. Crystallographic data for Limonin and its derivatives.

Molecular Formula	Limonin		Limonin acetic acid solvate		Limonin methanol solvate		epi-Limonin idoacetate		epi-Limonin chloroacetate	
	$C_{26}H_{30}O_8$		$C_{26}H_{30}O_8 \cdot CH_3COOH$		$C_{26}H_{30}O_8 \cdot CH_3OH$		$C_{26}H_{31}O_8(COCH_2I)$		$C_{26}H_{31}O_8(COCH_2Cl)$	
a (A)	14.52 ± 0.02 (14.53)		12.73 ± 0.02 (12.75)		12.89 ± 0.02		15.03 ± 0.02		12.26 ± 0.02	
b (A)	17.74 ± 0.02 (17.75)		16.79 ± 0.02 (16.85)		16.59 ± 0.02		12.36 ± 0.02		10.92 ± 0.02	
c (A)	8.87 ± 0.02 ( 8.88)		12.42 ± 0.02 (12.40)		12.19 ± 0.02		15.93 ± 0.02		11.50 ± 0.02	
$\beta$	90° 0'		90° 0'		90° 0'		95° 12' ± 15'		93° 15' ± 15'	
U (A <sup>3</sup> )	2284.8		2654.6		2606.8		2952.7		1836.4	
Z	4		4		4		4		2	
Dm (gm.cm <sup>-3</sup> )	1.366		1.313		1.260		1.426		1.220	
Dx (gm.cm <sup>-3</sup> )	1.366		1.315		1.280		1.441		1.228	
$\lambda$ (cm <sup>-1</sup> )	0		1.2		1.2		1.2		2.0	
$\lambda = 0.7107 \text{ \AA}$										
Mol. wt.	470.5	4	530.6	4	502.5	4	640.5	2	568.0	2
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub> - D <sub>2</sub>		P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub> - D <sub>2</sub>		P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub> - D <sub>2</sub>		P2 <sub>1</sub> - C <sub>2</sub>		P2 <sub>1</sub> - C <sub>2</sub>	



precisely the situation for a centrosymmetric electron density distribution. A Fourier synthesis using these phase-angles would provide an approximation to the electron density superposed on its spurious inversion image. The problem of unravelling the two images would be the same as that in the analysis of cholesteryl iodide by Carlisle and Crowfoot (1945), where extensive use of the chemically determined structure had to be made. While the problem of determining the atomic positions for the two molecules of epi-limonol iodoacetate in the asymmetric crystal unit would be initially formidable the final concurrence of the structures of two chemically identical but crystallographically distinct molecules would be compelling confirmation of the validity of the structure announced.

#### 2.4. The collection and processing of the experimental data.

##### 2.4.1. The collection of intensity data for crystals of epi-limonol iodoacetate. The epi-limonol

iodoacetate crystals obtained from aqueous acetone solutions were thin fragile plates with the (001) faces well-developed. Although occasionally forms such as (111) were observed, the plates usually had irregular boundaries.

For the purpose of X-ray photography rectangular parallelipeds of dimensions 1.0 x 0.2 x 0.02 mms. were cut from the irregular crystals so that the two longest sides were parallel to the optically determined extinction axes. In general it was not possible to predetermine whither the artificial needle axis so created was the a or b crystallographic

axis. This had to be determined by oscillation photographs taken about the needle axis (unlike those about the  $a$ -axis,  $b$ -axis photographs showed  $m$  symmetry since  $b$  represents the two-fold screw-axis).

It was observed that the normally stable iodoacetate crystals on exposure to X-rays became yellow, then brown, and finally disintegrated after a few days. Once this decomposition had been initiated it apparently continued even in the absence of X-radiation. Since however the lifetime of a crystal was fairly long, the only serious disadvantage was the necessity to cut and correctly orientate a succession of crystal specimens.

The intensity data was recorded using two moving-film techniques: successive levels about both the  $a$  and  $b$  axes were photographed using an equi-inclination Weissenberg goniometer, while those about the  $c$ -axis were obtained using a Buerger precession goniometer. For the first method a crystal was correctly orientated on goniometer arcs using the stationary crystal procedure of Weisz & Cole 1948. The arcs bearing crystal were then transferred to a Weissenberg goniometer of the equi-inclination type where it can be arranged that the direct X-ray beam and the diffracted beam are equally inclined to the reciprocal lattice layer being photographed. The value of the equi-inclination angle ( $\mu_0$ ) was determinable from the unit cell parameters, and from it, and charts appropriate to the goniometer in question, the screen setting necessary to isolate the zone of reflexions being recorded was obtained and verified by a taking a  $5^\circ$

oscillation photograph. Since the film-holder could only traverse a distance equivalent to some  $200^\circ$  rotation of the crystal specimen and for upper level  $\underline{a}$ -axis photographs  $(nkl)$  would therefore allow the recording of only  $(nkl)$  and  $(nk\bar{l})$  or  $(nkl)$  and  $(n\bar{k}l)$  reflexions and since for this space group  $|F(n\bar{k}l)| = |F(nkl)| + |F(nk\bar{l})|$ , it was desirable to arrange that the former set were photographed by adjusting the starting point for crystal rotation. The intensities were recorded on a pack of five films to allow correlation of the strongest and weakest reflexions. At the beginning of the investigation Cu  $K_\alpha$  radiation from a Philip's PW 1010 X-ray generating set run at 30 kilovolts and 12 milliamperes was used, but later all the data was obtained using Mo  $K_\alpha$  radiation. In this case the films in the multiple pack were interleaved with sheets of nickel foil to increase the absorption of the Mo  $K_\alpha$  radiation between successive films to a degree similar to that for Cu  $K_\alpha$  X-rays. In this way sixteen five-pack film series -  $(n_1kl)$ ,  $n_1 = 0 \dots 6$  and  $(hn_2l)$ ,  $n_2 = 0 \dots 8$  - were obtained. The remaining film series of type  $(hkn_3)$ ,  $n_3 = 0 \dots 2$ , were obtained using the versatile precession method (Buerger, 1944) in which diffraction is caused to occur by an unusual motion of a single crystal in a monochromatic X-ray beam. The crystal is supported at the unmoved point of a universal joint and a rational direction of the crystal is caused to take a precessing motion. In this way the diffraction spectrum is recorded as an undistorted scaled photograph of the reciprocal

crystal lattice. Apart from the convenience of such an undistorted picture with its resulting ease of indexing reflexions, the shape of the diffraction spots remains practically constant at all levels and facilitates visual estimation of intensities. The principal disadvantages of the precession method are that in comparison with the Weissenberg method it is capable of furnishing a record of only a limited part of the reciprocal lattice and that in photographs of the upper levels the record in the centre of the film is characteristically missing. These objections were not of consequence in the present investigation since such a wealth of intensity data was available from the Weissenberg series. Moreover there was a particular advantage in recording  $c$ -axis data by a method which did not require (as Weissenberg recording of the  $(hkn_3)$  zones would have) mounting the fragile, platy crystals on supporting fibres perpendicular to the plane of the plate, but only required crystals with the same setting as for the Weissenberg-recorded series. Precession camera techniques require that the crystal-to-film distance is maintained accurately constant and therefore it is impossible to use multiple pack series. Instead satisfactory correlation of the strong and weak reflexions is achieved by multiple exposure series - a series of films with the same record but different (and known) exposure times.

#### 2.4.2. The estimation and correction of intensities.

The scaling factors for reflexions on different films of

each of the nineteen series could be predetermined: for the Weissenberg film series the factors were obtained from a plot of the scaling factor as a function of the equi-inclination angle (Rossmann 1956) and for the precession film series they were predetermined by the exposure times chosen. The intensities of the diffraction spots were estimated visually using either the 'standard spot' technique, or by comparison with a series of calibrated spots. The first technique involves the choice of a spot of suitable intensity on one film of a series as a standard which immediately establishes the relative intensities of the corresponding spots on the other films of the series; using these reference spots the intensities of the other reflexions the series can be assigned. In the second technique an external standard series of spots is prepared by recording a crystal reflexion several times but with different exposure times. It has the advantages that the number of operations required to match a reference and unknown reflexion is less and that an arithmetic series of calibrated spots can be prepared so providing more accurate matching than can the geometric series of standard intensities derived in the 'standard spot' method.

The intensities of the reflexions for each series were measured twice and by a different operator on each occasion using different standards. The two estimations were placed on the same scale and compared. When individual

estimates differed by more than 15% further independent estimations were made when, in general, the required agreement was obtained. Each reflexion was then assigned an intensity equal to the arithmetic mean of the two approximately concurring estimations.

The intensity for an X-ray reflexion corrected for extinction, absorption, geometrical factors (Tunell 1939) and the effects of polarisation of the X-ray beam is proportional to the square of the structural amplitude for the crystal plane in question. The X-radiation (Mo K $\alpha$ ) and crystal dimensions had been chosen so as to minimise absorption and this effect and that of extinction were assumed to be negligible and no corrections were made for them. The polarisation correcting factor  $2(1 + \cos^2 2\theta)^{-1}$  is normally combined with the geometrical correcting factor which also varies from reflexion to reflexion, and depends on the relative time during which the conditions for reflexion remain satisfied. It is consequently inversely proportional to the velocity of passing through these conditions i.e. (for Weissenberg-recorded reflexions) inversely proportional to  $(\sin \theta)^{-1} (\cos^2 \mu_e - \cos^2 \theta)^{\frac{1}{2}}$  where  $\theta$  is the angle made by the reflecting crystal plane and the reflected beam, and  $\mu_e$  is the equi-inclination angle (Tunell, loc. cit.) Waser's analysis (1951 a, b) of the motion of the precession goniometer revealed that its velocity was dependent on the actual position in space of the precessing crystallographic axis. The values of

the correcting function (polarisation factor included) for 0.05 reciprocal lattice unit intervals of  $\Sigma$ , have been represented by contours on charts published by Grenville-Wells and Abrahams (1951). The charts are transparent and when the one appropriate to the level being considered is superposed on a scaled replica of the corresponding reciprocal net it is possible by visual interpolation to read off the correcting factor to be applied to each reflexion.

Each measured intensity of each plan's reflexion was corrected to give a relative value for the square of the structural amplitude from which the structural amplitude itself was derived.

2.4.3. Determination of film-to-film scaling constants. There now existed experimental values of some three thousand independent structure amplitudes ( $|F_0|$ 's) derived from the moving film series  $(n_1 \ kl)$ ,  $n_1 = 0 \dots 6$ ,  $(h \ n_2 \ l)$ ,  $n_2 = 0 \dots 8$ ,  $(hkn_3)$ ,  $n_3 = 0 \dots 2$ . In general the sets of data from each series were on different scales and the problem was to place the observed results on a uniform, if arbitrary, scale. The procedure adopted involved the choice of the  $(0kl)$  series as the arbitrary standard. The  $(h \ n_2 \ l)$  series were then correlated using the appropriate members of the  $(n_1 \ kl)$  group in this wise:  $\sum |F(0 \ n_2 \ l)|$  for each  $n_2 = 0 \dots 8$  was determined from the  $(h \ n_2 \ l)$  series and  $\sum |F(0kl)|$  for each  $k = 0 \dots 8$  from the  $(0kl)$  series data, the summations including for each  $k = n_2$  only those  $|F|$ 's common to the

(0kl) and (h n<sub>2</sub> l) zones. Every  $K(0 n_2 l) = \frac{\sum IF(0kl) I}{\sum IF(0 n_2 l) I}$   $k = n_2 = 0 \dots 8$  was determined and was a factor suitable for placing the IFI's derived from the (h n<sub>2</sub> l) series on the same scale as those from the (0kl) series. There was also determined  $R(0) = \frac{\sum_k \sum_l IF(0kl) I}{\sum_{n_2} \sum_l IF(0 n_2 l) I}$ . Repetition of this procedure using the (h n<sub>2</sub> l) sets of IFI's and successive (n<sub>1</sub> kl) sets determined  $K'(n_1 n_2 l) = \frac{\sum IF(n_1 kl) I}{\sum IF(n_1 n_2 l) I}$   $k = n_2 = 0 \dots 8$  and  $R(n_1) = \frac{\sum_k \sum_l IF(n_1 kl) I}{\sum_k \sum_l IF(n_1 n_2 l) I}$  and  $K(n_1 n_2 l) = R(0) \cdot K'(n_1 n_2 l) / R(n_1)$ , so that again the K's related the data to the (0kl) standard. In this way seven (n<sub>1</sub> = 0 ... 6) scaling factors  $K(n_1 n_2 l)$  were obtained for each n<sub>2</sub> = 0 ... 8 and  $\bar{K}(h n_2 l) = \frac{1}{7} \sum_{n_1} K(n_1 n_2 l)$  was chosen as the scaling factor for each (h n<sub>2</sub> l) set.

Similarly scaling factors to place each (n<sub>1</sub> kl) set of data on the same scale as the (hol) set were derived. These factors  $\bar{K}'(n_1 kl)$  were used to derive factors  $\bar{K}(n_1 kl) = \bar{K}'(n_1 kl) / \bar{K}'(0kl)$  consistent with the  $\bar{K}(h n_2 l)$  group.

The c-axis data (h k n<sub>3</sub>) was similarly correlated with the rest, using (h n<sub>2</sub> l) data. More direct scaling using (n<sub>1</sub>kl) data was not undertaken because of the greater quantity of (h n<sub>2</sub> l) data available.

It was then possible to place every structural amplitude determined on the same scale and to average the values obtained from more than one series.

It is certain that the values of  $\bar{K}$  obtained are no



more than good approximations to the best values which could be derived by minimising the 'residual'  $E = \sum_i \sum_j (K_i J_{ij} - K_j J_{ji})^2$  where  $K_i$  is the scaling constant for the  $i$ -th film and  $J_{ij}$  is the sum of all reflexions on film  $i$  common to film  $j$  and  $J_{ii} \equiv 0$ . An equivalent result could be obtained by refining the constants using the equations  $K_i = \sum_j K' J_{ij} J_{ji} / \sum_j J_{ij}^2$  where the primed  $K$ 's are the initial values and the new values are normalised after each cycle by division by some  $K_i$  - say  $K(Okl)$ . (Dickerson, 1959).

The procedure described by Dickerson where the ratio of the sums of the intensities of all corresponding reflexions were used implicitly emphasised the contributions of the strong reflexions in contrast to that described by Kraut (1958) where use of the average values of the ratios of the intensities of corresponding reflexions gave greater weight to the weak reflexions. In the present work, using the ratio of the sums of structure amplitudes introduced an effective weight of  $I^{-\frac{1}{2}}$  to Dickerson's method and thereby gave satisfactory weight to all reflexions.

It was not considered necessary to improve the values of the  $\bar{K}$ 's at that stage since it was expected that more satisfactory scaling could eventually be carried out by relating the structure amplitudes for each zone to those calculated when the structure reached an advanced stage of refinement.

## 2.5. The determination of the structure of limonin.

At various stages in the long process of recording, estimating and correcting the intensity data the observed structure amplitudes available were used in assaults on the structural problem. The information gleaned using first the incomplete data and finally the complete experimental data is described in this section.

2.5.1. The two-dimensional investigation. At an early stage structural amplitudes derived from intensities for the (hol) and (okl) zones were available. With the squares of the (hol) structure amplitudes as coefficients a Patterson Fourier series was computed to give the projection of the atomic vector distribution perpendicular to the screw axis of the crystal. No sharpening modifications of the Patterson function were made and the vector distribution (fig. 2.1) was used to assign  $\underline{x}$  - &  $\underline{z}$  - co-ordinates to the two iodine atoms in the crystal asymmetric unit. Using these co-ordinates structure factors for the (hol) zone were calculated assuming a mean temperature factor of 4. Since the b - (screw) axis projection of the electron density distribution is centred the phase-angles for the (hol) zone are either zero or  $\pi$  and the great relative scattering power of iodine atoms would result in some 80% of the phase-angles calculated being correct. The calculated phase-angles were combined with the observed structure amplitudes in a Fourier calculation of a substantially correct electron density distribution projected on (010). Fig. 2.3a. shows such a distribution subsequently calculated using Mo  $K_{\alpha}$ - derived data, a temperature factor of 4.94 and phase-angles from

more accurate iodine co-ordinates. This projection perpendicular to an axis some 12 Å long did not (as indeed it was not expected to) reveal any structural features, although a useful refinement of the iodine atom co-ordinates resulted. Fig. 2.3b shows the later-determined atomic positions projected on (010) and it is noteworthy that carbon atoms, 6, 9, 16, 28, 29, 3<sup>1</sup>, 16<sup>1</sup>, 22<sup>1</sup>, 28<sup>1</sup> and oxygen atoms 2, 7, 8, 2<sup>1</sup>, 7<sup>1</sup> are well resolved in Fig. 2.3a and carbons 17, 23, 13<sup>1</sup>, 20<sup>1</sup> and oxygens 5 and 3<sup>1</sup> partially so.

In a similar fashion the (100) Patterson vector distribution yielded  $\underline{y}$  - and  $\underline{z}$  - co-ordinates for the iodine atoms (Fig. 2.2). Phase-angles calculated from these co-ordinates were used in a Fourier synthesis of the (100) electron density distribution. The non-centred nature of this projection meant that the phase-angles could possess any value between zero and  $2\pi$  so that the (100) projection is inherently less accurate than the (010). As Fig. 2.4 shows this non-centred projection down a 16 Å axis is relatively featureless and its only utility was another refinement of the iodine atom co-ordinates.

2.5.2. The Harker Section at  $y = \frac{1}{2}$ . When a crystal has a twofold screw axis (b), vectors between symmetry-related atoms fall on section  $y = \frac{1}{2}$  of the Patterson function. Such a section is known as the Harker section (Harker, 1936). When reflexion data for levels  $(n_1 k l)$ ,  $n_1 = 0, 1$ ,  $(h n_2 l)$ ,  $n_2 = 0 \dots 7$ , and  $(hko)$  had been collected and processed,

computation of the Harker section at  $y = \frac{1}{2}$ ,  $P(x \frac{1}{2} z) = V^{-1} \sum \sum \sum (-1)^k |F(hkl)|^2 \cos 2\pi (hx + lz)$ , was undertaken to provide more accurate co-ordinates for the iodine atom positions. Some 2,000 coefficients ( $|F(hkl)|^2$ 's) were included in the series. To reduce the negative regions  $F^2(000)$  was arbitrarily assigned a value of 400. The symmetry of the vector distribution required only half of it to be computed and it was found convenient to calculate its value at intervals of  $a/30$  in the  $\underline{a}$ -axis direction and  $c/30$  in the  $\underline{c}$ -axis direction. When the Harker peaks had been located the value of the function in the region containing the peaks was recomputed at intervals of  $1/60$  of the unit cell edge in both axial directions, and accurate co-ordinates for the peaks determined by graphical interpolation. Fig. 2.5a shows the Harker peaks, representing vectors between iodine atoms in the unit cell related by the screw-axis, and a psuedo-Harker peak, representing a vector between non-symmetry-related iodine atoms whose  $y$  - co-ordinates differ by  $\frac{1}{2}$ . Fig. 2.6. shows section  $y = \frac{1}{2}$  of the later-calculated Patterson function and reveals that the inclusion of a further 1,000  $|F|^2$  terms has resulted in a decline in the prominence of the psuedo-Harker peak. From the work described in subsections 2.5.1 and 2 the co-ordinates in Table 2.2 were assigned to the two independent iodine atoms and were not varied in the subsequent analysis until a late stage of the refinement of the structure.

Table 2.2. Co-ordinates of iodine atoms

	x	y	z
I	0.0617	0.0367	0.0292
I	0.1333	0.0367	0.3658

### 2.5.3. The three-dimensional investigation.

Using a 'Mercury' electronic digital computer through the agency of Dr. O.S. Mills (Manchester University), structure factors (SFI) were calculated, on the basis of the iodine co-ordinates in Table 2.2, for those planes for which structural amplitudes had been determined from the observed data. The reliability index,  $R = | |F_o| - |F_c| | / |F_c|$ , ( $|F_o|$  and  $|F_c|$  are respectively the observed and calculated structure amplitudes, the former being on the same scale as the latter) was found to have the low value of 0.35 in keeping both with the non-centrosymmetric space group and the dominance of the iodine atoms as X-ray scatterers. The phase-angle approximations from SFI and the 2,950 observed structure amplitudes ( $|F_o|$ 's) weighted according to the method of Sim (1959) were used to calculate the corresponding three-dimensional approximation of the electron-density distribution (FI). The weighting of the structural amplitudes permitted the inclusion at this early stage of all the structural amplitude data rather than merely a part as is normally the case with other more subjective assessments of the reliability of phase-angle information. The symmetry element in the space group of epi-limonol iodoacetate required that the electron-density function was computed for only half the unit cell volume and in that volume the function was evaluated at intervals of 1/48 of the cell edge in all three principal directions. A master grid with these intervals on a scale of 2 cms to the Angstrom

was constructed on paper for sections of the unit cell perpendicular to the screw axis and from this master copies were made. The calculated values for the electron density were inserted at the appropriate grid-point on the correct section sheet. Over each section thus prepared, a perspex sheet was placed and on this a contoured representation of the electron density was drawn in wax pencil. The perspex sheets were then stacked and the resulting model examined for recognisable molecular features. In F1 the iodine atoms appeared in the same sites ascribed to them by the previous work but were surrounded by successive shells of negative electron density. In addition, the number of prominent maxima which might correspond to the positions of the carbon and oxygen atoms in the molecule was some three times greater than the number of atoms being sought, and no arrangements of the maxima into chemically acceptable features were apparent. At this stage it was feared that diffraction effects introduced by the dominant scatters (the iodine atoms) and the possible additional false symmetry resulting from the similarity of the iodine atoms' y-co-ordinates were responsible for the obscurity of F1, and that solution of the molecular structure problem would be unusually difficult.

Two projects to help to unravel F1 were then undertaken. Since the (010) projection of the electron density distribution was centred, phase-angles for the (h0l) zone could only take values of zero or  $\pi$  in contrast to the

variable values between zero and  $2\pi$  allowed by a non-centric distribution such as occurs in the three-dimensional case, and since the relative scattering power of the iodine atoms was so great, more than half of the phase-angles assigned to the (hol) zone by considering only iodine contributions would be correct, while half the remainder would be correct by chance. It could therefore be claimed that the projected electron density distribution shown in Fig. 2.3a was inherently more correct than the three-dimensional distribution even if it was less directly useful because of the impossibility of its resolution. For these reasons it was considered that maxima in FI which corresponded to maxima in the projection were more likely to represent true atomic sites than those which did not, and it was hoped that inclusion of the former in a phase-angle calculation would give phasing accurate enough to produce successful interpretation of further Fourier syntheses. Some 40 atomic positions were chosen by comparing FI and the projection, and several structure factor calculations for the (hol) zone, made using the sets of co-ordinates in different groups of twenty. (The calculations, like all subsequent ones, were made on the Glasgow University DEUCE computer). In no case did the agreement index (R), which for this zone was 0.48 when only iodine parameters were used, fall below 0.43.

The further possibilities of this procedure were not explored since other lines of attack were showing more signs of success, but retrospective examination showed that somewhat more than half of the positions chosen corresponded to real atomic sites.

The second method of attack involved the computation of the three-dimensional sharpened Patterson function and the laborious operation of transferring the origin to each of the iodine atoms in turn, and then the derivation from the resulting vector distributions of a minimum function (see Sub-Section 1.3.2f). The minimum function derived ( $P_{\min}$ ) possessed a number of maxima of the same order as the number of unknown atomic positions.

At this time it was discovered that a subtle programming error in the Manchester computation of F1 had introduced the large number of spurious maxima, and a corrected Fourier synthesis (F2) of the electron density distribution revealed an arrangement of maxima readily associated with a plausible chemical partial structure. The accurate concurrence of the majority of the maxima in  $P_{\min}$  and F2 allowed the confident assignment of co-ordinates to some fifty eight atoms. The co-ordinates were chosen by visual interpolation from the contours on a perspex sheet representation of F2 and consequently had an accuracy no greater than 0.01 of a cell edge. These co-ordinates were included, with the original iodine co-ordinates, in a structure factor calculation (SF2) where an arbitrary mean temperature factor (B) of 4.94 was used, and all the light atoms were assumed to be carbons. Unsatisfactory scaling of the meant that no value for R was obtained but the phase-angles, which are independent of such considerations were included with the (unweighted) observed structure amplitudes in a new Fourier synthesis (F3), from which co-ordinates for all the



light atoms were obtained. The co-ordinates obtained by visual interpolation were included in a structure factor calculation ( $B = 4.94$ ) where it was found that  $R = 0.26$ . A further structure factor calculation (SF3) using more accurate co-ordinates obtained by the method of Booth (1948) but excluding the contribution of C17 and the atoms of the furan ring attached to it gave  $R = 0.24$ . The excluded atoms were so treated because imperfect resolution made confident allocation of accurate co-ordinates difficult. The relative heights of the maxima in F3 allowed use of the correct form factors for all the atoms included in the calculation except O9 which was still treated as a carbon; O3 and O3<sup>1</sup> did not enjoy greater peak height in F3 but were required to be oxygens by the chemical evidence (private communication from D.H.R. Barton).

2.5.4. The refinement of the structure. The reluctance to define O9 as an oxygen rather than a carbon atom and the exclusion of six atoms from our calculations was not due to any ignorance of their kind and position but resulted from a desire to assume only that which was certain and not merely that which was highly probable. Therefore it can be said that, at this stage, the structure of *epi*-limonol (see Fig. 2.7) and hence of limonin was established and that the succeeding operations were required only to improve the agreement of the observed data with the structure assigned.

The phase-angles from SF3 were included in a new Fourier series (F4) and the refined co-ordinates of the light atoms determined by Booth's method. It was considered

necessary to represent the electron-density function by contours only in the region of the six atoms which had still to be assigned accurate co-ordinates. Consideration of F4 still did not allow these co-ordinates to be chosen accurately.

It was now hoped to accelerate the structural refinement by the method of Least Squares (Hughes 1941; Booth 1947) in the form provided by a standard N.P.L. DEUCE programme. This programme furnishes structure factors (for the co-ordinates provided), the shifts to be applied to atomic co-ordinates, and the optimum anisotropic temperature parameters. Unfortunately it was designed only to accept problems involving 64 atoms, and O1, 2, 3, 4, 5, 6, 7, 8, 9 were excluded in addition to Cl7 and its attached furan ring. Even then it was not found possible to solve the normal equations from some of the sets of data and never possible to solve any for the iodine parameters. R was not evaluated but new co-ordinates were derived for 62 light atoms and it was decided to resume refinement by the previous 'F<sub>o</sub>' synthesis' method. The later trend of the refinement makes it most doubtful that utilisation of the least squares programme was wise at this stage.

The new co-ordinates provided by the least squares calculations, the original iodine co-ordinates and the F4 co-ordinates for O1, ..., 9 were used to calculate structure factors (SF5). B was given a value of 4.66 and R was found to be 0.22. Fourier series (F5) using the phase-angles from

TABLE A.

Relationship between co-ordinates  
listed in Table 2,3 and atoms in  
Fig. (E) etc.

~~The atom numbering is the revised version described in  
the addendum.~~

For unlisted atoms the identity operation  $(x, y, z)$  has  
to be assumed.

<u>Atom</u>	<u>Operation</u>	<u>Atom</u>	<u>Operation</u>
C 3	$1-x, \frac{1}{2} + y - 1, -z$	C 5'	$2-x, \frac{1}{2} + y - 1, 1-z$
4	$x, y, z-1$	6'	$2-x, \frac{1}{2} + y - 1, 1-z$
19	$1-x, \frac{1}{2} + y, 1-z$	7'	$1+x, y, z$
21	$1-x, \frac{1}{2} + y, 1-z$	9'	$1-x, \frac{1}{2} + y - 1, 1-z$
28	$x, y, z-1$	10'	$2-x, \frac{1}{2} + y - 1, 1-z$
29	$1-x, \frac{1}{2} + y - 1, -z$	19'	$1+x, y, z$
0 1	$x, y, z-1$	28'	$2-x, \frac{1}{2} + y - 1, 1-z$
2	$1-x, \frac{1}{2} + y - 1, -z$	29'	$2-x, \frac{1}{2} + y - 1, 1-z$
3	$1-x, \frac{1}{2} + y, 1-z$	30'	$1+x, y, z$
9	$1-x, \frac{1}{2} + y, 1-z$	31'	$1+x, y, z$
I'	$1+x, y, z$	32'	$1+x, y, z$
C 1'	$1-x, \frac{1}{2} + y - 1, 1-z$	0 1'	$2-x, \frac{1}{2} + y - 1, 1-z$
2'	$1-x, \frac{1}{2} + y - 1, 1-z$	2'	$2-x, \frac{1}{2} + y - 1, 1-z$
3'	$2-x, \frac{1}{2} + y - 1, 1-z$	3'	$2-x, \frac{1}{2} + y - 1, 1-z$
4'	$2-x, \frac{1}{2} + y - 1, 1-z$	4'	$1+x, y, z$
		5'	$1+x, y, 1+z$

**TABLE 2.3.** Atomic co-ordinates for epilimonol iodoacetate at each stage of refinement after F3. (See Text).  
The atoms are numbered as in Formula (I).<sup>(k,4)</sup> Primed co-ordinates distinguish the two crystallographically independent molecules. Some peak heights at the later stages are included.

Atom	Peak Heights	X	Y	Z	Peak Heights	X	Y	Z
I		.13498	.03490	.36600	.06220	.46150	.97214	
		.13592	.03331	.36858	.06219	.45951	.97281	
	C1	.54752	.14088	.01350	.02796	.36452	.32802	
		.54448	.13991	.01462	.02906	.36455	.32995	
		.54350	.13054	.01231	.03499	.35022	.32682	
		.54195	.13982	.01030	.03306	.35857	.32572	
		.54302	.14274	.00981	.03118	.36388	.32724	
	C2	.63450		.00042	.03535	.34059	.42931	
		.63440	.10230	.00042	.03695	.33294	.43104	
.63585		.09365	.00467	.04864	.32676	.43222		
.63323		.09849	.00410	.04176	.33082	.43216		
.63359		.09884	.00285	.03906	.33336	.43195		
C3	.37500	.47917	.01475	.93952	.35344	.44960		
	.37579	.48358	.00898	.93952	.35347	.45391		
	.37434	.48249	.00562	.94314	.36697	.45078		
	.37409	.48027	.00946	.94154	.36140	.45072		
	.37615	.48148	.01112	.94141	.35840	.45151		
C4	.39835	.10158	.94104	.90969	.33146	.22342		
	.40541	.11120	.94220	.90765	.33771	.22074		
	.40937	.11007	.93618	.91038	.34050	.22008		
	.40906	.10609	.93810	.90915	.33582	.22215		
	.40756	.10353	.93756	.90927	.33321	.22348		

Original co-ordinates from Table 2.2 used.

717  
811

772  
754

661  
692

776  
780

692  
695

690  
684

670  
656

733  
761

TABLE 2.3. (CONTINUED)

<u>Atom</u>	<u>Peak Heights</u>	<u>X</u>	<u>Y</u>	<u>Z</u>	<u>Peak Heights</u>	<u>X</u>	<u>Y</u>	<u>Z</u>
C5		.38494	.13636	.03862		.95587	.45561	.22146
		.38848	.12905	.03739		.95476	.44718	.21505
		.39827	.13957	.03332		.94023	.44134	.21915
	704	.39610	.13261	.03243	707	.95110	.44104	.21661
	694	.39515	.13030	.03412	751	.95353	.44126	.21688
C6		.31806	.10310	.10508		.89825	.53850	.16973
		.31618	.10597	.08844		.89966	.53832	.16643
		.30847	.10314	.08438		.89651	.53855	.17165
	737	.31272	.10632	.08167	763	.89818	.53962	.17151
	767	.31407	.10484	.08125	734	.89860	.53667	.17153
C7		.31869	.15702	.16527		.02294	.12800	.83629
		.31834	.15772	.16418		.02712	.13001	.83105
		.30902	.15778	.17132		.03247	.12426	.82973
	684	.31914	.15699	.16645	763	.02933	.12730	.83382
	743	.32123	.15951	.16414	786	.02718	.12732	.83521
C8		.41750	.14340	.22324		.96396	.15311	.76694
		.41262	.13630	.21145		.96918	.16970	.76030
		.41410	.13838	.23019		.96153	.16545	.75579
	773	.41283	.13845	.21649	733	.95872	.16970	.75321
	790	.41281	.13734	.21482	729	.95741	.16362	.75442
C9		.49419	.16667	.16346		.07479	.54628	.29444
		.49186	.16139	.16299		.07731	.54859	.29086
		.49032	.16862	.16319		.07930	.55671	.29539
	729	.49067	.15989	.16366	823	.07380	.55440	.29465
	714	.48940	.16489	.16399	850	.07254	.55223	.29453

TABLE 2.3. (Continued)

Atom	Peak Heights	X	Y	Z	Peak Heights	X	Y	Z
C10		.48631	.10094	.07604		.99673	.47215	.30637
		.48586	.10524	.07680		.00480	.46977	.30853
		.48231	.09799	.07737		.00103	.46887	.30505
	686	.48665	.10290	.07650	765	.00023	.47268	.30620
	691	.48830	.10378	.07676	825	.99876	.47368	.30564
C11		.57633	.17248	.21858		.87500	.11204	.63331
		.57751	.18076	.22415		.87467	.10484	.63869
		.58624	.15954	.22038		.86615	.08516	.64118
	673	.57751	.18076	.22415	771	.86057	.08465	.63221
	581	.58143	.18070	.22049	849	.86001	.08034	.63013
C12		.58337	.27211	.28533		.79167	.16271	.66060
		.57782	.27158	.28697		.79112	.16622	.65764
		.58500	.26797	.29245		.79701	.17101	.64907
	701	.58324	.26821	.28627	805	.79612	.16622	.64985
	653	.58097	.27048	.28091	788	.79467	.16555	.65025
C13		.49529	.30779	.30735		.78314	.17977	.66494
		.50086	.30344	.30712		.78481	.18697	.74621
		.48908	.29715	.30137		.79308	.18122	.75181
	656	.49423	.30402	.30247	-	.78481	.18679	.74621
	689	.49523	.30602	.30351	852	.78460	.17445	.74603
C14		.43125	.22213	.29577		.88369	.21465	.79508
		.42481	.22346	.29412		.88052	.22181	.78921
		.42425	.22953	.29485		.87379	.21635	.78585
	808	.42624	.22323	.29566	776	.87803	.22181	.78689
	749	.42885	.21685	.29704	781	.87955	.21412	.78825

TABLE 2.3. (Continued)

Atom	Peak Heights	X	Y	Z	Peak Heights	X	Y	Z
C15		.35769	.21584	.36102		.87181	.26515	.87431
		.35576	.21493	.36273		.86980	.25558	.87286
		.35323	.21267	.35917		.86317	.25259	.87514
	640	.35296	.21493	.36080	705	.86917	.26006	.87346
	634	.35226	.21506	.36140	739	.86984	.26279	.87336
C16		.35448	.31344	.41106		.78514	.28686	.89942
		.35224	.30192	.41027		.79050	.28690	.90356
		.34644	.32264	.41401		.79181	.28266	.88632
	679	.34936	.31853	.41081	710	.78775	.28350	.89344
	670	.35007	.31755	.41012	676	.78631	.28206	.89581
C17		.50686	.32456	.40540		.73239	.27869	.75946
						.72663	.27602	.75950
		.50948	.35144	.39296	724	.72670	.28130	.76637
	716	.51113	.34834	.39662	761	.72357	.28206	.75788
						.72247	.28052	.75650
C18		.44323	.39127	.24856		.74498	.08706	.78321
		.44316	.39130	.24852		.74396	.09517	.79071
		.44605	.39999	.25184		.74405	.09927	.79657
	684	.44451	.39807	.25319	680	.74657	.09257	.79044
	709	.44482	.39376	.25154	694	.74690	.08942	.78752
C19		.49452	.47111	.91450		.06333	.02146	.62167
		.48596	.47518	.91198		.06399	.02711	.62527
		.49381	.48266	.90794		.06305	.03629	.63276
	749	.49098	.47921	.91296	684	.06324	.03248	.62569
	745	.49030	.47700	.91508	730	.06993	.02773	.62330

\*Corrigendum: The coordinates listed for the circled atoms should refer to the uncircled ones.

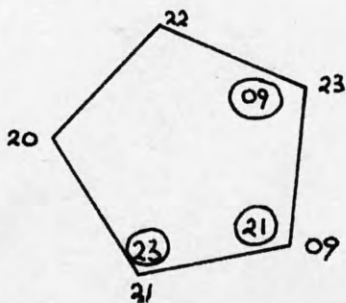




TABLE 2.3. (Continued)

<u>Atom</u>	<u>Peak Heights</u>	X	Y	Z	<u>Peak Heights</u>	X	Y	Z
C20		.52546	.35694	.41579		.62319	.25786	.70967
						.62598	.26571	.71009
		.57680	.43428	.41806	801	.62678	.26262	.71958
	687	.57033	.43472	.41688	821	.62423	.26465	.70980
		.62983	.33111	.44792 *		.62361	.26519	.70970
C21						.55592	.23131	.75152
						.54708	.22796	.75113
		.37759	.10247	.55602 *	686	.54523	.22779	.74768
	726	.37600	.09915	.55679 *	679	.55008	.23360	.75044
		.54389	.51746	.42129		.55310	.23627	.75307
C22						.58721	.29384	.62652
						.59176	.29500	.63146
		.67925	.43162	.41789	643	.58543	.30113	.62543
	680	.67716	.43358	.41745	689	.59159	.29504	.62773
		.68510	.03652	.42517 *		.59407	.29323	.62911
C28						.50879	.31359	.62329
						.51381	.31125	.62600
		.45266	.00965	.57957 *	706	.51404	.30968	.62166
	640	.45568	.01252	.57939 *	699	.51769	.30569	.62323
		.35235	.18921	.87844		.51823	.30055	.62246
C28		.35309	.19045	.87946		.82096	.35729	.23958
		.35151	.19458	.87930		.81935	.33800	.23821
		.35309	.19045	.87946		.81177	.33460	.23770
	707	.35192	.18758	.87963		.81089	.33786	.24015
						.81281	.33905	.24075

TABLE 2.3. (Continued.)

<u>Atom</u>	<u>Peak Heights</u>	<u>X</u>	<u>Y</u>	<u>Z</u>	<u>Peak Heights</u>	<u>X</u>	<u>Y</u>	<u>Z</u>
C29		.63352	.49410	.09394		.91777	.27548	.14544
		.63432	.48460	.08025		.91739	.27390	.15039
		.62390	.47798	.08688		.90643	.27832	.14341
	760	.62642	.48389	.08490	814	.91281	.27465	.14027
	836	.62755	.48481	.08463	787	.91481	.27286	.14246
		.41229	.03094	.25235		.02756	.22537	.69850
		.41352	.03279	.25023		.01641	.23558	.69710
		.41732	.02641	.25376		.01904	.24207	.70935
	760	.41414	.02867	.25221	693	.99880	.24298	.70874
	695	.41420	.02923	.25167	638	.00829	.24456	.70553
C31		.17312	.16844	.21954		.09937	.24342	.94335
		.17228	.15866	.22019		.09536	.24287	.94464
		.17877	.14923	.21634		.09787	.23798	.95323
	634	.17470	.15866	.21699	722	.09807	.23506	.95020
	696	.17256	.16143	.21607	789	.09851	.23427	.94991
		.10192	.10154	.23617		.14548	.33771	.97114
		.09865	.09965	.23369		.14835	.33290	.97332
		.09403	.10438	.23443		.13753	.33708	.96656
	742	.09794	.10522	.23804	577	.14439	.33515	.96959
	717	.10103	.10476	.23988	581	.14535	.33296	.97051
O1		.49933	.12038	.93000		.96114	.28811	.29225
		.47594	.12407	.93225		.96089	.28625	.29529
		.48951	.12124	.93006	1217	.96641	.28819	.29666
	1031	.49657	.12130	.93127	1305	.96280	.28704	.29383
	491					.96028	.28719	.29289

TABLE 2.3. (Continued)  
Peak Heights

<u>Atom</u>	<u>Peak Heights</u>	X	Y	Z	<u>Peak Heights</u>	X	Y	Z
02		.33935	.41769	.05256		.91279	.29450	.49948
		.33786	.42487	.05410		.91294	.29375	.49807
	1020	.33847	.42172	.05351	1027	.91378	.28635	.49910
	972	.33901	.42316	.05564	1002	.91210	.28765	.49982
		.45696	.42052	.97296		.91079	.28800	.49971
03		.45328	.42116	.97255		.89412	.43750	.41667
		.44919	.41869	.97495		.89490	.44316	.41852
	1086	.45209	.41953	.97589	1090	.89731	.43567	.42149
	1132	.45268	.42043	.97784	1111	.89477	.43754	.41810
		.25273	.11323	.20308		.89254	.43737	.41716
04		.25058	.11761	.20458		.07108	.22231	.87104
		.25065	.11536	.20613	1021	.07130	.22193	.87184
	1124	.25023	.11474	.20720	1084	.07268	.22659	.86928
		.15181	.25344	.19594		.07120	.22758	.87024
		.16166	.26771	.19482		.07022	.22637	.87061
05		.16359	.26773	.19399	883	.09042	.16121	.00060
	935	.16452	.26588	.19389	1034	.08722	.16599	.00134
		.44010	.14894	.37806		.09001	.11953	.00367
		.43558	.14635	.37809		.08765	.14016	.00373
		.43764	.14682	.37844	1009	.08701	.14957	.00312
06	1017	.43803	.14459	.37867	1002			
						.88875	.33217	.80079
						.88915	.32719	.79960
						.88972	.32927	.79684
						.88771	.33084	.79748
					.88697	.33236	.79868	

TABLE 2.3. (Continued)

<u>Atom</u>	<u>Peak Heights</u>	X	Y	Z	<u>Peak Heights</u>	X	Y	Z
07		.28229	.33590	.44731		.76906	.30250	.97796
		.27581	.33002	.44126		.76796	.30011	.97549
	1031	.27550	.33342	.43651	1099	.77033	.30209	.99705
	1030	.27643	.33501	.43453	1131	.76992	.29963	.97695
						.76913	.29723	.97618
08		.41994	.37511	.43387		.70444	.30256	.84310
		.41987	.37513	.43381		.71097	.30119	.83891
		.41760	.37212	.42575		.71309	.29963	.84120
	1037	.42066	.37345	.42737	1079	.70970	.30236	.83951
	1049	.42106	.37303	.42749	1104	.71418	.30294	.83749
09		.69521	.40729	.41667 *		.47856	.25096	.69275
						.47861	.25523	.69453
		.68309	.52902	.43899 *	804	.47750	.25042	.69353
		.68830	<del>.47008</del>	.43894 *	991	.47786	.25192	.69595
			<del>.46646</del>			.47230	.25183	.69681
	557	.53354						

(SF5) allowed co-ordinates to be allotted to all atoms in the two molecules in the asymmetric crystal unit. At and from this stage shifts in the iodine co-ordinates were also determined. O9' was confirmed to be an oxygen atom but it was not certain that the newly introduced O9 was so.

The whole procedure was reiterated in SF6 (B = 4.66, R = 0.198) and F6, followed by SF7 (B = 4.94, R = 0.197) and F7 which confirmed O9 as an oxygen atom. The progress of the refinement is outlined in Table 2.3.

It is hoped that in the interval since this was written final refinement by the elimination of series termination errors by an  $F_c$  synthesis (Booth 1946, 1947b), by improvement of the temperature factor(s) used, and perhaps by more satisfactory correlation of the original sets of intensity data, will have been achieved.

## 2.6. The nature of the structure.

The investigation just described revealed limonin to be a highly oxygenated triterpenoid compound with the functional groups previously reported. The complete stereochemistry was determined and is displayed in I and Fig. 2.8. It was possible to define not merely the positions of all the atoms in the two independent molecules in the asymmetric crystal unit, but also to describe the atomic species of each atom by considering relative peak heights in the successive Fourier syntheses.

At the point where the X-ray work was able to provide a definitive solution to the constitutional and stereochemical problem, partial structure (IV) had been unambiguously established by the more classical methods of organic analysis (D.H.R. Barton, private communication 1959). Since that time, Professor Barton and his associates have been able to rationalise the great mass of chemical experimental data on limonin by using the structure described above. (Arrigoni et. al., 1959).

That borohydride reduction of limonin had indeed yielded the equatorially configured epimeric limonol was confirmed. The ring junctions A/A shows cis, A /B trans, B/C trans and C/D trans fusion, while the angular methyl groups (C 18 and C 30) have axial conformations. In the decalin part of the molecular skeleton rings B and C have 'chair' and 'boat' conformations respectively.

Biogenetically limonin is a bicarbocyclic triterpenoid related to euphol (V) Barton et al. 1954, 1955; Arrigoni et al., 1954) from which the four atoms terminating the methyl-isoheptenyl sidechain have been removed and the remainder converted into a furan ring. Ring A has been oxidatively cleaved between C 3 and C 4 and the C 3 carbonyl thus formed oxidatively cyclised on to C 19 - there are twin precedents for this C 3 - C 4 biogenetic cleavage in a triterpenoid skeleton in the recently revealed structures of dammarenolic acid (VI) and nycanthic acid (VII) (Arrigoni etc. al., 1959, Whitham, 1959). The limonin

structure would further require a Baeyer-Villiger oxidative cleavage of a ring D 16 - ketone, followed by recyclisation to give the ring D lactone; one methyl group would require to migrate from C 14 to C 8. Since oxidation of dihydrobutyrospermol (VIII) by chromic acid yields a 7 - ketone (IX) (Lawrie et al., 1956), it is not inconceivable that further oxidation and Baeyer-Villiger reaction would yield successively the 16 - ketone, the epoxidised ketone, and finally an epoxy-lactone of the same partial structure as limonin - see (X) - (XII).

The numbering convention for the atoms in limonin follows from these biogenetic hypotheses. Where it has been necessary to distinguish atoms of the two molecules in the asymmetric crystal unit primed numbers have been used.

## 2.7. Conclusion.

Thus the 127 years old problem of the structure of limonin has been solved, and in the process it has been shown that a 70 - atom structural problem (two chemical molecules were concerned, hydrogens were ignored) can be successfully elucidated by the methods of X-ray analysis, using no more chemical information than a molecular formula, and the knowledge that an alcoholic hydroxyl function could be produced in each molecule. It can probably be asserted that this later information would have been unnecessary, and the crystallographic problem no more difficult in essence, had the crystallographically suitable solvates which were sought been available.

2.8. REFERENCES

- ARNOTT, S., and ROBERTSON, J.M., (1959), *Acta Cryst.*, 12, 75.
- ARRIGONI et al., (1954), *Helv. chim. Acta*, 37, 2303.
- ARRIGONI et al., (1959), *Experientia* (to be published).  
 et al., (1954), *Chem. and Ind.*, 1325.  
 et al., (1955), *J.C.S.*, 876.
- BERNAY, (1841), *Annalen*, 40, 317.
- BOOTH, A.D., (1946), *Proc. Roy. Soc. A*, 188, 77.  
 (1947a), *Nature*, 160, 196.  
 (1947b), *Proc. Roy. Soc. A*, 190, 482.  
 (1948), 'Fourier Technique in Organic Structure Analysis'  
 P. 64. (C.U.P.)
- BUERGER, M.J. (1944), 'Photography of the Reciprocal Lattice.'  
 (ASKRED monograph No. I).
- CARLISLE, C.H., and CROWFOOT, D., (1945), *Proc. Roy. Soc. A*, 184, 6.
- CHANDLER, B.V., and KEFFORD, J.F. (1951), *Aust. J. Sci.*, 14, 24.
- COREY, E.J., et al., (1958), *J.A.C.S.*, 80, 1204.
- DICKERSON, R.E., (1959), *Acta Cryst.*, 12, 610.
- EMERSON, O.H., (1948), *J.A.C.S.*, 70, 545.  
 (1952), *ibid*, 74, 688.
- FUJITA, A., and WADA (1931), *J. Pharm. Soc. Japan*, 51, 506.
- FUJITA, A., and HIROSE, Y. (1954), *ibid*, 74, 365.  
 (1956), *ibid*, 76, 129.
- GEISSMAN, T.A., and TULAGIN, V. (1946), *J. org. chem.*, 11, 760.
- GRENVILLE-WELLS, J., and ABRAHAMS, S.C. (1951), *Rev. Sci. Inst.*, 23, 328.
- HUGHES, E.H., (1941), *J.A.C.S.*, 63, 66.
- JONES, F.T., and PALMER, K., (1949), *J.A.C.S.*, 71, 1935.
- KATMATSU (1902), *J. Pharm. Soc. Japan*, 24, 979.
- KEFFORD, J.F., and CHANDLER, B.V., (1951), *Aust. J. Sci.*, 14, 55.



- KOLLER, G., and CZERNY, H., (1936), *Mh. Chem.*, 67, 248.
- KRAUT, J., (1958), *Acta Cryst.*, 11, 895.
- KUBOTA, T., and TOKOROYAMA, T., (1957), *Chem. & Ind.*, 1928.
- LAWRIE, et al., (1956), *J. org. Chem.*, 21, 491.
- MELETA, et al., (1957), *Helv. chim. Acta*, 40, 1420.
- ROSENFELD, R.S., and HOFFMANN, K., (1951), *J.A.C.S.*, 73, 2491.
- ROSSMANN, M.G., (1956), *Acta Cryst.*, 9, 819.
- SHICHTER, M.S., and HALLER, H.L., (1940), *J.A.C.S.*, 62, 1307.
- THOMS, (1923), *Ber. deut. Pharm. Geo.*, 33, 68.
- TUNELL, G., (1939), *Amer. Min.*, 24, 448.
- WHITHAM, G., (1959), *Proc. chem. Soc.*, 271.
- WINCHELL (1943), 'Optical Properties of Organic Compounds.'  
(U. of Wisconsin Press.)

## CHAPTER III.

The preliminary investigation of the structure of the macrolide antibiotics.3.1. Introduction

The term macrolide encompasses a group of therapeutically and biogenetically important antibiotics characterised by a lactone moiety of large ring size. Methymycin (XIII) was the first antibiotic of this class for which a complete structure could be presented (Donin et al., 1954; Djerassi et al., 1956a, b, c, d). Methymycin is produced in a fermentation process by a *Streptomyces* strain (Djerassi et al., 1956a), and from the mother-liquors after its removal the isomeric neomethymycin (XIV) can be isolated. The determination of the constitution of this second macrolide has been described by Djerassi (1957, 1958a, b). A further isomer pykromycin also produced by a *Streptomyces* strain can, on the basis of reported degradations, (Anliker, 1957a, b; Brockmann 1957), only have structure (XIII) (Epimeric at C - 3 with methymycin) or (XV). Other antibiotics in this series whose structures have been determined include magnamycin (XVI) (Woodward 1957) and erythromycin (XVII) (Wiley, 1957).

The structural elucidation of the antibiotics of this group is not only of intrinsic interest but as the number of secure structure determinations increases, the statistical validity of the biogenetic speculations of

Woodward (1957) and Birch (1957) is enhanced. Further observations of biogenetic significance can be made by the assignation of absolute configuration to the asymmetric centres in these macrolides and some progress towards this end has been reported (Djerassi, 1958b).

### 3.2. The X-ray crystallographic investigation of the macrolide antibiotics.

X-ray work in this field has been undertaken with the threefold objective of resolving the constitutional ambiguity in pykromycin, of establishing the configuration at each asymmetric centre in methymycin, neomethymycin, and pykromycin, and of defining the conformation of these macrocyclic structures.

3.2.1. Macrolides and macrolide derivatives. The three macrolides named above are isomeric and have molecular formula  $C_{25}H_{43}N_7O_7$ . Samples of each were available, that of neomethymycin in the form of its methylene dichloride solvate. Each macrolide contains the amino-sugar group, of desosamine, and for methymycin and neomethymycin the aglycones (sugar-free residues) methynolide and neomethynolide were also available. It was felt that the successful achievement of structure determination in this field would lie, as with limonin, in the investigation of suitable heavy atom derivatives. Neomethymycin methylene dichloride solvate was one such derivative and it was decided to undertake a crystallographic study of this compound simultaneously with attempts to prepare others of a similar nature.

The chloroacetic esters of the aglycones and a quaternary methiodide of pykromycin have been made but it has not yet been found possible to grow single crystals of them suitable for X-ray diffraction studies. In addition the neomethymycin methylene dichloride solvate has been desolvated by heating to 100°C at 5 mm pressure for 24 hours. Crystallisation of the neomethymycin so obtained from chloroform and carbon tetrachloride solutions has yielded solvates which are at present being investigated.

3.2.2. Crystallographic data for pykromycin and neomethymycin methylene dichloride solvate. Pykromycin and the methylene dichloride solvate of neomethymycin both crystallise from methylene dichloride in the orthorhombic system, the former with a prismatic habit and the latter in six-sided plates.

Zero-layer precession camera photographs taken with Mo  $K_{\alpha}$  X-radiation yielded the unit cell parameters shown in Table 3.1. In both cases the sole diffraction conditions were that (hoo), (oko), and (ool) were present only for h, k, l even so that the space group was  $P2_1 2_1 2_1 - D_2^4$ . Crystal densities were determined in the usual way and showed that for both substances the unit cell contained four molecular formula weights.

3.2.3. Neomethymycin methylene dichloride solvate. Further study of this solvate is being undertaken at the same time as other derivatives are being prepared and examined.

In neomethymycin methylene dichloride solvate the

TABLE 3.1.Crystallographic data for neomethymycin and pykromycin.

	<u>Neomethymycin methylene dichloride solvate</u>	<u>Pykromycin</u>
Molecular formula	$C_{25}H_{43}NO_7 \cdot CH_2Cl_2$	$C_{25}H_{43}NO_7$
$a$ (Å)	$12.42 \pm 0.02$	$11.26 \pm 0.02$
$b$ (Å)	$20.92 \pm 0.02$	$25.71 \pm 0.02$
$c$ (Å)	$11.57 \pm 0.02$	$10.48 \pm 0.02$
$V$ (Å <sup>3</sup> )	3008.8	3036.6
$Z$	4	4
$D_m$ (gm. cm. <sup>-3</sup> )	1.19	1.01
$D_x$ (gm. cm. <sup>-3</sup> )	1.22	1.03
Molecular weight	554.4	469.5
Space group	$P2_1^2 2_1^2 - D_2^4$	$P2_1^2 2_1^2 - D_2^4$

scattering power of the chlorine atoms is by no means dominant,  $z_{Cl}^2 / \sum z_L^2 = 0.4$ , so that a further purpose was added to our investigation, viz., the possible demonstration of the success of a structural determination where the 'heavy' atom was rather less dominant than usual.

Multiple exposure series for five levels (0 ... 4) along the c-axis and for the (hol) zone were obtained using the precession camera. Multiple pack series for five levels along the a-axis were obtained using an equi-inclination Weissenberg goniometer. Mo  $K_{\alpha}$  radiation was used in each case. The intensities of the reflections were estimated by comparison with a series of calibrated spots. The estimations were done in duplicate using different standards, and such was the quality of the reflexions recorded that for the Weissenberg series the discrepancy between successive estimations never exceeded 12% and for the precession series they did not exceed 8%. The experimental data was processed and correlated in the manner described for the epi-limonol iodoacetate data yielding a group of some 1,700 independent structural amplitudes. This number is rather less than might have been expected and is apparently caused by the rather large temperature factor which results in a rapid diminution of intensity of the reflections as the scattering angle is increased.

Projections of the Patterson vector distribution down all three principal crystallographic axes were computed on DEUCE both with and without sharpening modification. A

sharpened three-dimensional Patterson distribution has also been computed using all the data. The three Harker sections of this last function i.e. at  $x = \frac{1}{2}$ ,  $y = \frac{1}{2}$  and  $z = \frac{1}{2}$  are shown in Figs. 3.1, 3.2, and 3.3. As these figures show, the chlorine-chlorine vectors are not sufficiently dominant for the maxima due to them to be immediately identified. At present the maxima of the Patterson functions (both two - and three - dimensional) are being investigated intensively in an attempt to determine the positions of the two chlorine atoms in the asymmetric crystal unit. Use is being made of the fact that the Cl - Cl distance in methylene dichloride is known. When the chlorine positions have been determined it is hoped to proceed to a structure elucidation by routes similar to those described in Chapter II.

S.S.      REFERENCES.

ANLIKER, R., and GUBLER, K., (1957a) *Helv. chim. Acta*, 40, 119.  
 (1957b) *ibid*, 40, 1768.

BROCKMANN, H., and OSTER, R., (1957) *Chem. Ber.*, 90, 605.

DJERASSI, C., and Coworkers

(1956a), *J.A.C.S.*, 78, 1729

(1956b), *ibid*, 78, 1733

(1956c), *ibid*, 78, 2907

(1956d), *ibid*, 78, 6390

(1957), *ibid*, 79, 2022

(1958a), *Tetrahedron*, 3, 255

(1958b) *ibid*, 4, 369

DOMIN, M.N., et. al., (1954), 'Antibiotics Annual,'  
 1953-4, pp. 179 - 185  
 (Med. Encycl. Inc. N.Y.)

WILEY, et al., (1957), *J.A.C.S.*, 79, 6062.

WOODWARD, R.B., (1957), *Angew. Chem.*, 69, 50.



## APPENDIX 1.

The computation of Fourier series.1. Introduction

Two of the aims of X-ray crystallographic structural studies viz. structure determination for large and complex molecules which have defied the classical methods of structural attack, and the structure determination of simpler molecules more rapidly and efficiently than by the ordinary techniques, require the aid of devices to reduce the labour and time involved in computation. These are rendered lengthy by the great wealth of data involved or by complexity of the calculus of the procedures or both.

The collection and estimation of intensities will normally be followed by their correction and scaling and by the determination of structural amplitudes. Patterson Fourier series may then be evaluated and structural features established. This will usually be succeeded by many structure factor calculations and Fourier syntheses involving the evaluation of many - termed Fourier series at many points in space (in the study of epi-limonol-iodoacetate some 3,000 terms were included and the function evaluated at  $49 \times 49 \times 25$  points) and perhaps least squares refinement cycles.

The recent introduction of a DEUCE general purpose digital computer in Glasgow has considerably expedited computation leaving the collection and estimation of intensities as the 'slow stage' in structure analysis. Input and output on DEUCE are in the form of punched cards, and it is now possible by means

of standard programs to proceed with the succession of correction and correlation of experimental data; and its use in various calculations after one initial punching of data cards bearing the index and intensity of each reflexion. Only a relatively small number of further parameters have to be punched and introduced during the subsequent analysis.

Since it was only at a late stage of the analysis of  $\alpha$ -limonol iodoacetate that high speed digital computer services were available, all the initial calculations had to be effected with only the aid of desk calculators. The evaluation of Fourier series was therefore one of the main preoccupations and was the motive for outlining the aid given by other devices in this appendix.

## 2. The evaluation of Fourier series.

### 2.1. The nature of the function to be evaluated.

The most general form for the Fourier series in three dimensions is  $f(x, y, z) = \sum_{h=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} F(hkl) \exp(-2\pi i (hx + ky + lz))$ . In general the Fourier coefficients are complex

$$F(hkl) = A(hkl) + i B(hkl)$$

The determination of the electron density distribution ( $\rho$ ) in a crystal is somewhat simplified by the fact that the function -

$$\rho(x, y, z) = 1/V \sum_{h=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} F(hkl) \exp(-2\pi i (hx + ky + lz)) \dots (1)$$

is real and therefore

$$F(hkl) = F(\bar{h} \bar{k} \bar{l})$$

$$\text{i.e. } A(hkl) = A(\bar{h} \bar{k} \bar{l}) \text{ and } B(hkl) = -B(\bar{h} \bar{k} \bar{l})$$

Equation (1) can also be simplified where the symmetry properties

of the space group of a crystal introduce relationships among classes of coefficients. For instance, in space group Pm, unique axis b, the electron density is given by

$$\rho(x, y, z) / 4 = \sum_0^{\infty} \sum_0^{\infty} \text{CCC} (hkl) \text{ccc} + \text{SCS} (hkl) \text{scs} + \\ \text{SCC} (hkl) \text{scc} + \text{CCS} (hkl) \text{ccs}$$

where CCC (hkl) etc. are respectively

$$\begin{aligned} & A (hkl) + A (\bar{h} \ kl) \\ - & A (hkl) + A (\bar{h} \ kl) \\ & B (hkl) - B (\bar{h} \ kl) \\ & B (hkl) + B (\bar{h} \ kl) \end{aligned}$$

and ccs etc. are  $(\cos 2\pi hx \cos 2\pi ky \sin 2\pi lz)$  etc.

$$\rho(x, y, z) / 4 = \sum_{k=0}^{\infty} C(x, y, k) \cos 2\pi ky \\ = \sum_{k=0}^{\infty} \left[ \sum_{l=0}^{\infty} \text{CC}(x, kl) \cos 2\pi lz + \right. \\ \left. \sum_{l=1}^{\infty} \text{CS}(x, kl) \sin 2\pi lz \right] \cos 2\pi ky$$

$$= \sum_{k=0}^{\infty} \left\{ \sum_{l=0}^{\infty} \left[ \sum_{h=0}^{\infty} \text{CCC} (hkl) \cos 2\pi hx \right. \right. +$$

$$\left. \sum_{h=1}^{\infty} \text{SCC} (hkl) \sin 2\pi hx \right] \cos 2\pi lz +$$

$$\left. \sum_{l=1}^{\infty} \left[ \sum_{h=0}^{\infty} \text{CCS}(hkl) \cos 2\pi hx + \sum_{h=1}^{\infty} \text{SCS}(hkl) \sin 2\pi hx \right] \sin 2\pi lz \right\}. (2)$$

This expansion has shown that the three-dimensional Fourier series can be reduced to a succession of simple operations involving only one-dimensional cosine and sine series, taking, in this case, the summations over h first, then over l, and finally over k. Greater efficiency in the calculations is achieved by use of the symmetry properties of these cosine and sine series.

In calculating the electron density or the complete asymmetric unit ( $X = 0$  to  $1$ ,  $Y = 0$  to  $\frac{1}{2}$ ,  $Z = 0$  to  $1$ ) in the space group  $Pm$  the following sequence of operations might be followed.

a) The determination from  $A(hkl)$  etc. of the values of the Fourier coefficients like  $CCC(hkl)$

b) The separation of the coefficients  $CCC(hkl)$  etc. into groups with index  $h$  even and  $h$  odd

c) The evaluation for  $X = 0$  to  $\frac{1}{4}$  at some interval of the eight summations like

$$\sum_{h=0}^{8} CCC(hkl) \cos 2\pi hx, \quad h \text{ even}$$

$$\sum_{h=1}^{8} CCC(hkl) \cos 2\pi hx, \quad h \text{ odd}$$

d) Appropriate combination of these odd and even summations to obtain  $CC(x, kl)$  and  $CS(x, kl)$  for  $X = 0$  to  $1$

e) Division of these new functions into groups with  $l$  even odd and the evaluation of the for  $z = 0$  to  $\frac{1}{4}$

four summations like  $\sum_{l=0}^{8} CC(x, kl) \cos 2\pi lz$

f) Combination of the new summations to give  $C(xz, k)$  for  $Z = 0$  to  $1$  and all  $X$

g) Evaluation for  $Y = 0$  to  $\frac{1}{4}$  and all  $x$  and all  $z$  of

$$\sum_{k=0}^{8} C(xy, k) \cos 2\pi ky, \quad k \text{ even}$$

$$\sum_{k=0}^{8} C(xz, k) \cos 2\pi ky, \quad k \text{ odd}$$

h) Combination of these results to give  $\sqrt{\rho(x,y,z)} / 4$  for all X, all Z, and  $Y = 0$  to  $\frac{1}{2}$ .

### 2.2.1. The Beevers-Lipson strip method of computation.

This device was first produced by Lipson & Beevers (1936) and its newer version was introduced by Beevers (1952). Values of the functions  $C \cos 2 \pi n h x$  and  $C \sin 2 \pi n h x$  are printed on two sets of strips, each strip carrying the values of the function for a particular value of the amplitude C and order h and all values of n. The newer strips have  $x = 1/120$  and  $n = 0 \dots 30$  with odd values on back of the strip and even values on the front. The functions are tabulated to the nearest whole number and C has values ranging from  $\pm 1$  to  $\pm 100$  at unit intervals and thence to  $\pm 900$  at intervals of a hundred.

For computation the strips for the appropriate C and h are selected from a box and the arguments summed for each nx using a desk adding machine. For work in 1/60ths it is only necessary to use the even values of n printed on the front of the strips.

An accuracy of the order of 0.2% can be obtained using the strips although this may be increased especially when large values of the amplitude C occur since two strips require to be used and the rounding-off errors may be combined.

### 2.2.2. The use of Robertson's Fourier Synthesiser (RUFUS)

A digital computer designed for Fourier synthesis has been described by Robertson (1954, 1955). It is a mechanical analogue of the strip method in which suitable gear ratios are employed to generate the sine and cosine functions, the outputs being recorded

by printing revolution counters. Backlash errors are rendered unimportant by the fact that only complete revolutions of the output shafts are counted and the gear ratios chosen to produce an error less than  $5 \times 10^{-8} \%$  i.e. five-figure accuracy.

The intervals chosen in the machine available in the Chemistry Department at Glasgow University were at  $6^\circ$  (1/60ths). A row of generators for each term corresponded to a set of Beevers-Lipson strips. The necessary amplitude C for each row can be applied by effecting C revolutions to the appropriate input shaft. A particular input shaft is engaged by a simple level and the power provided by a reversing electric motor capable of operating at high speeds so that a coefficient may be set in a few seconds.

### 2.2.3. The use of the DEUCE electronic digital computer.

Much use was made of a Fourier synthesis program provided by the National Physical Laboratory. The data used are separated cards bearing index and the real (A) and imaginary (B) part of  $F_o$  (or  $F_c$ , or  $F_o - F_c$ , or  $|F_o|^2$ ). The data is indexed so that the number of different h values is least, and the number of k values. next lowest. The data cards (punched in decimal) were used as input behind a binary converter programme whose variables included a) the combination of indices hkl which should be tested to determine the class of each reflexion, i.e. the trigonometrical functions for each reflexion, b) the contribution of A and B to the series for each class of reflexion (the possible contributions are any or all of  $A_{ccc}$ ,  $-A_{css}$ ,  $-A_{scs}$ ,  $-A_{ssc}$ ,  $B_{ccc}$ ,  $B_{csc}$ ,  $B_{ccs}$ ,  $-B_{sss}$ ), c) an indication of whether reflexions with  $k < 0$  and/or  $l < 0$  were included, d) the volume (V) of the unit cell, and

e) a suitable scaling factor (K). The binary output obtained by use of the converter programme is placed behind a 'section card' which gives the limiting values of  $x$ ,  $y$ ,  $z$  for the region required and the intervals in each direction between successive summation points. This pack read into the computer behind the main Fourier programme. The output cards contain eight decimally punched function values for constant  $y$  and eight equally spaced values of  $x$ . The computation of  $F_0$  syntheses for limonin in which some 3,000 terms were involved took about fifteen hours when the intervals of summation were 1/48ths in all directions and where summation covered half the unit cell.

#### REFERENCES

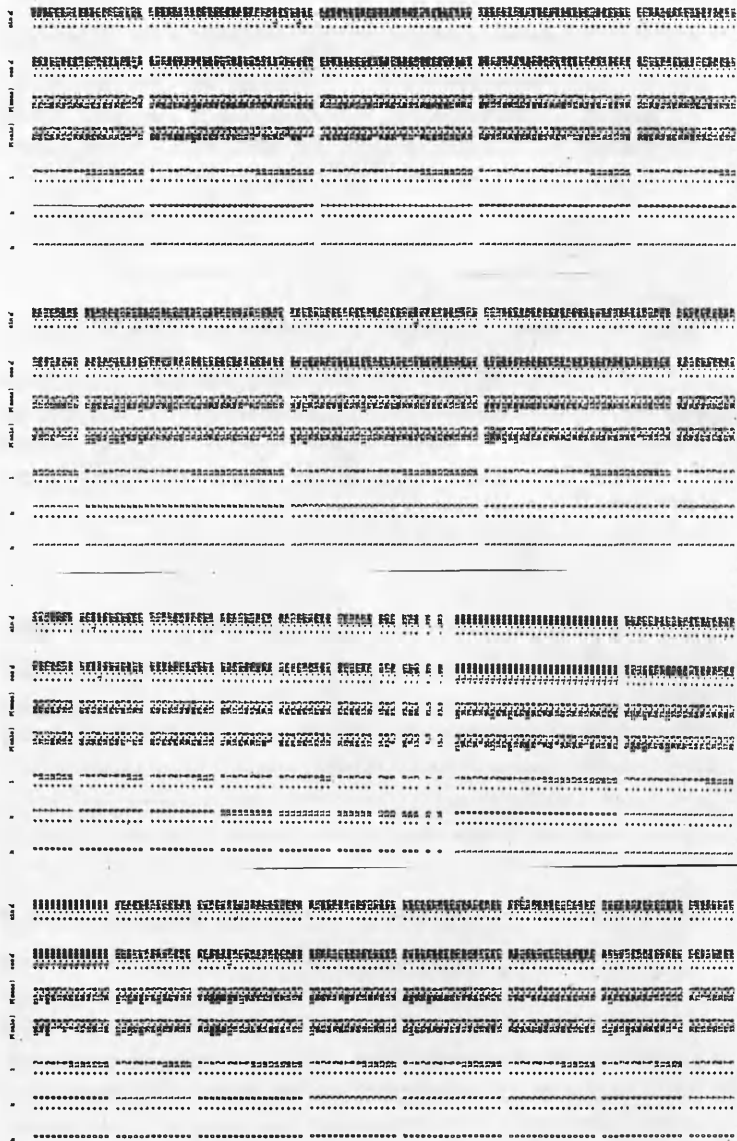
- BEEVERS, C.A., (1952), *Acta. Cryst.*, 5, 670.  
LIPSON, H., and BEEVERS, C.A., (1936), *Phil Mag.*, 17, 855.  
ROBERTSON, J.M. (1954), *Acta Cryst.*, 7, 817.  
(1955), *Acta Cryst.*, 8, 286.

APPENDIX II.

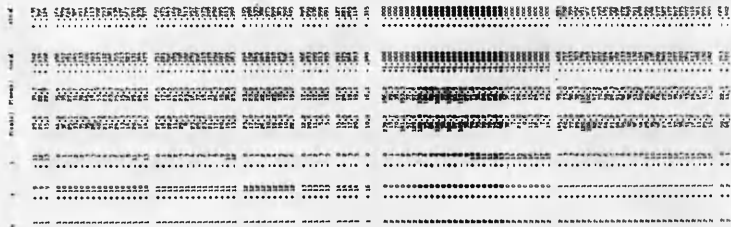
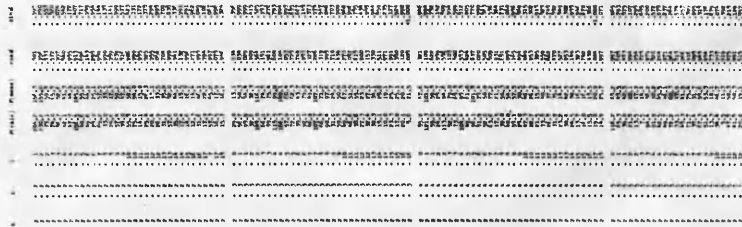
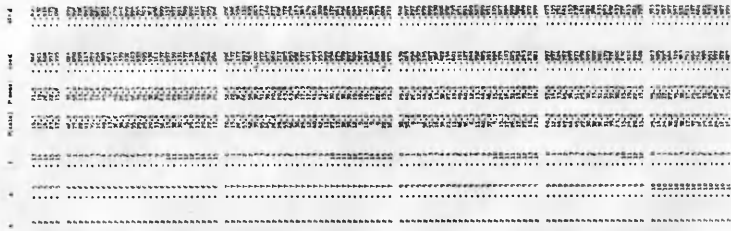
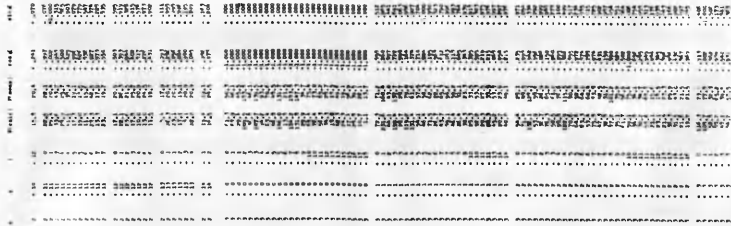
In this appendix appears a comparison of the structural amplitudes for epi-limonol iodoacetate calculated using the last set of co-ordinates in Table 2.3 with those which were observed experimentally. The phase-angle for each reflexion is given implicitly by the values of  $\cos \alpha$  and  $\sin \alpha$ . At this stage of the refinement the reliability index was 0.19 - 0.20.



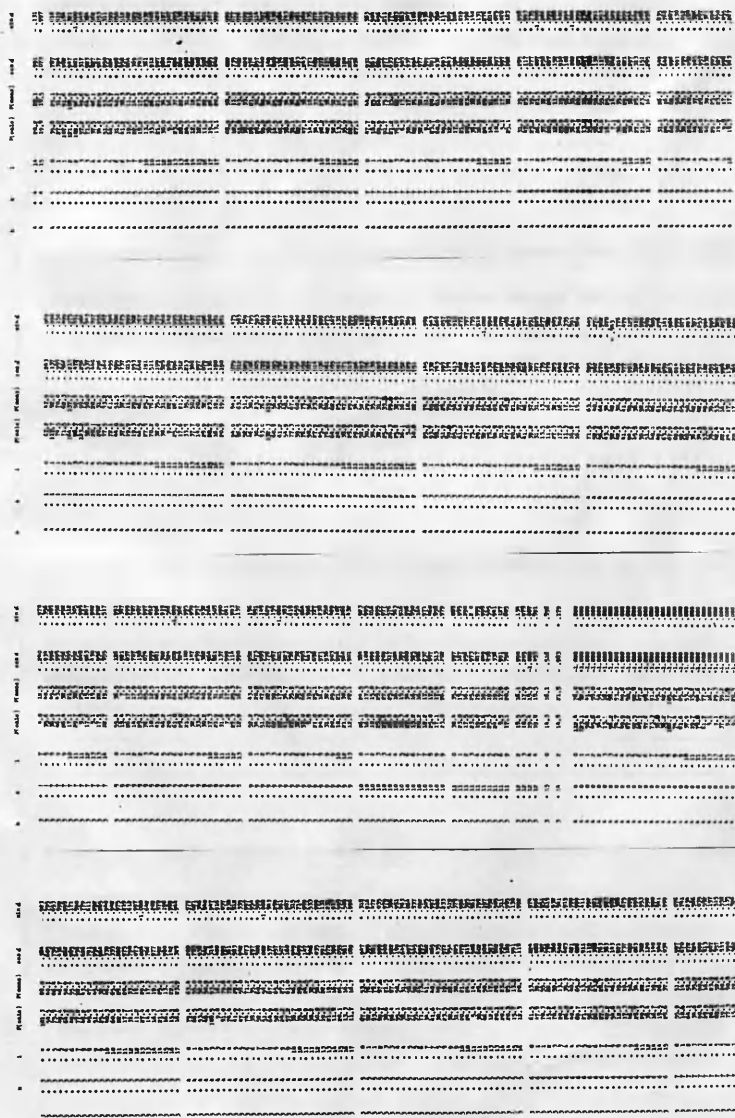
Epi-limonol iodacetate :  $F(\text{calc.})$ ,  $F(\text{meas.})$  and phase-angles.



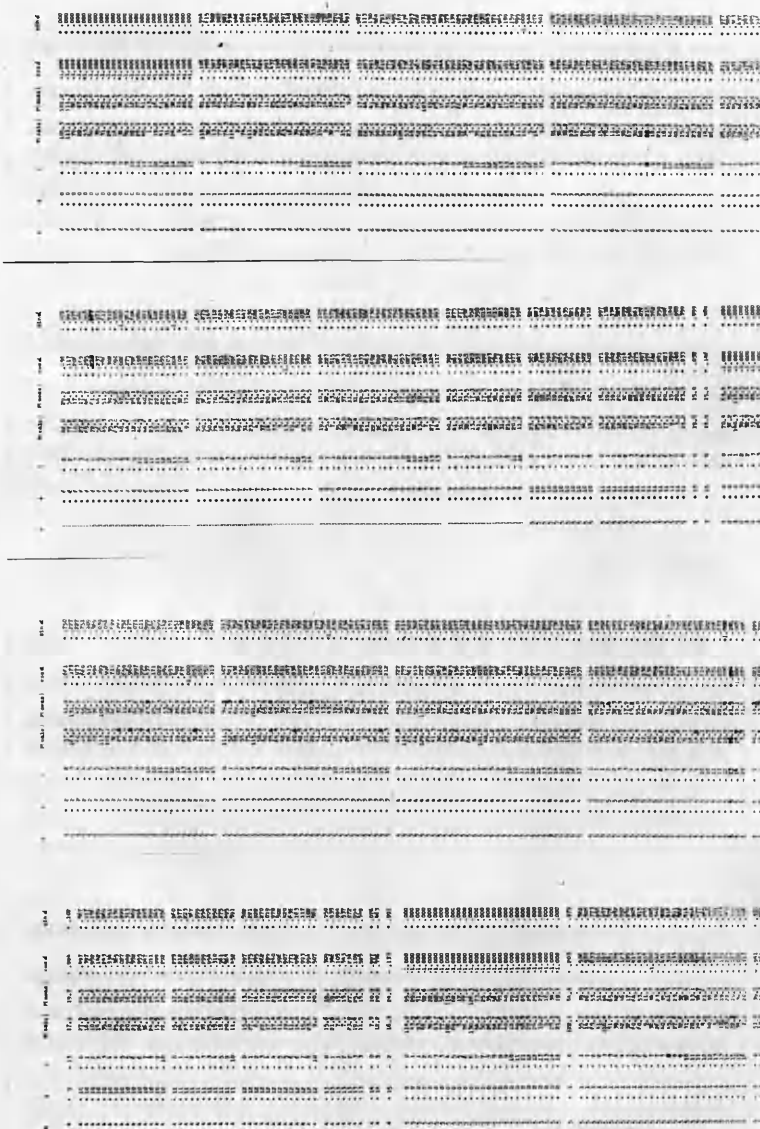
Epi-limonol iodacetate : F(calc), F(meas) and phase-angles.



*Epi-limonol iodacetate* :  $F(\text{calc.})$ ,  $F(\text{mens})$  and  $\text{phase-angles}$ .



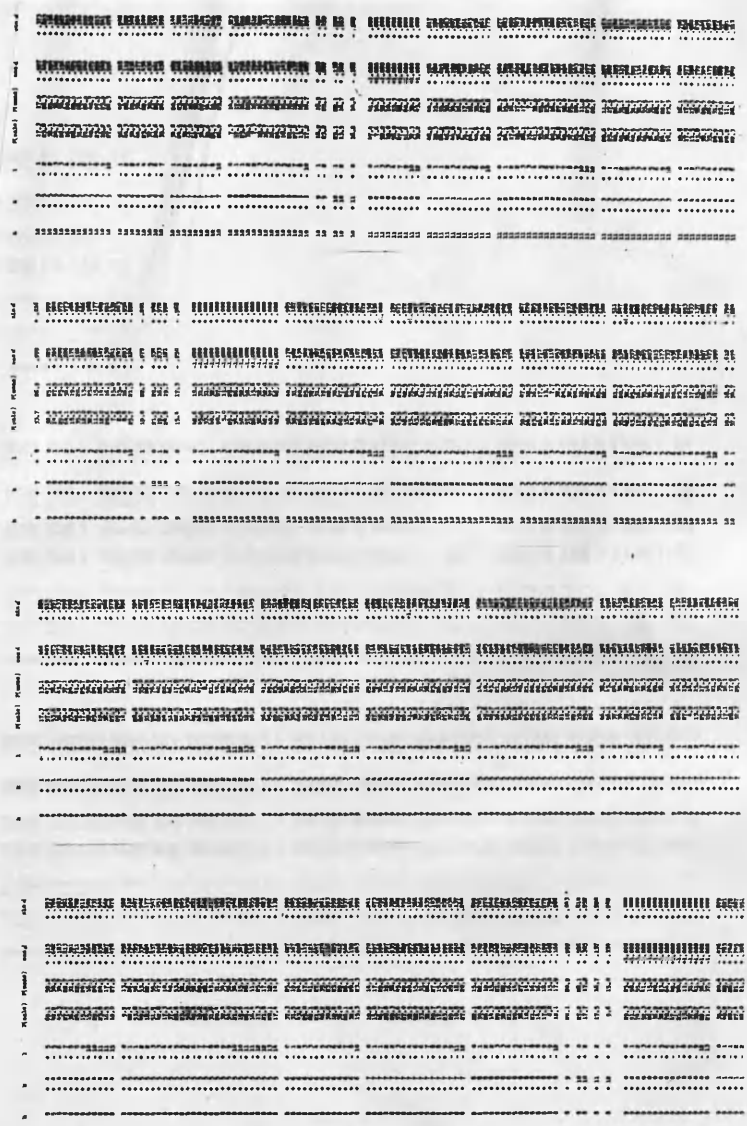
*Epi-limonol iodacetate* : *F(calc)*, *F(meas)* and *phase-angles*.



Epi-limonol iodoacetate :  $F(\text{calc.})$ ,  $F(\text{meas.})$  and phase-angles.



*Epi-limonol iodacetate : F(calc), F(meas) and phase-angles.*





APPENDIX III.

Diagrams and Illustrations.



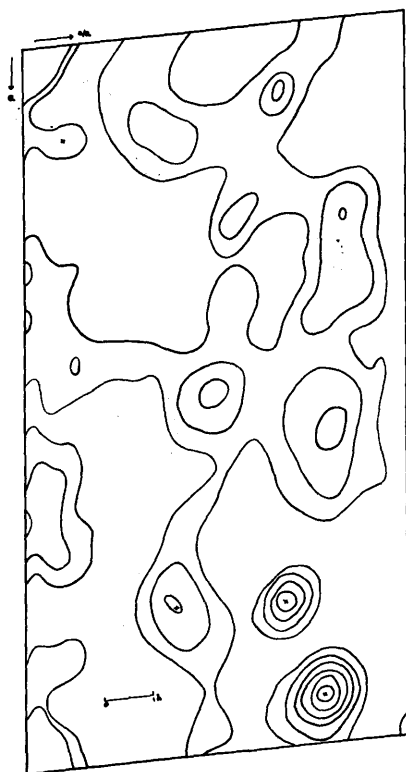
## \* Illustrations of the structure of epi-limonol iodoacetate

It should be emphasised that the structure depicted in Figs 2.3b, 2.7, 2.8 is what it was conceived to be at the earliest stages of our knowledge.

As soon as it was possible to assign positions to the atoms, a working model of the asymmetric crystal unit and its contents was created by erecting, on a cork base, balls on wires showing the disposition of the atoms. Fig. 2.8. is a view of this model at the beginning of the refinement. By using this model as a template for Fig. 2.3b., two errors were perpetuated

(i) C 32 was wrongly assumed to be bounded to the iodine atom appearing in .b-axis projection at  $x = 0.93781$ ,  $z = 1.02719$  instead of (correctly) to its symmetry analogue at  $x = 1.06219$ ,  $y = 0.45951$ ,  $z = 0.97281$ .

(ii)  $x$  (C2'),  $x$  (C1'),  $z$  (C1) were given the wrong sign.



*Fig. 2.1. Patterson vector distribution projected on (010) with the iodine-iodine vectors shown bearing a cross.*

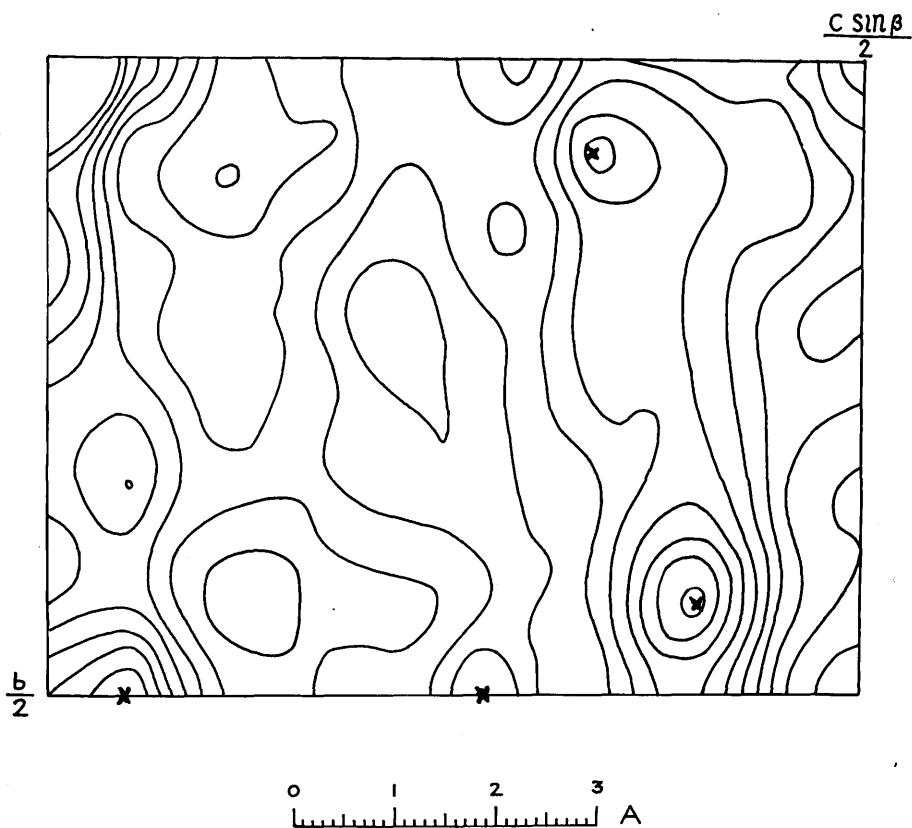
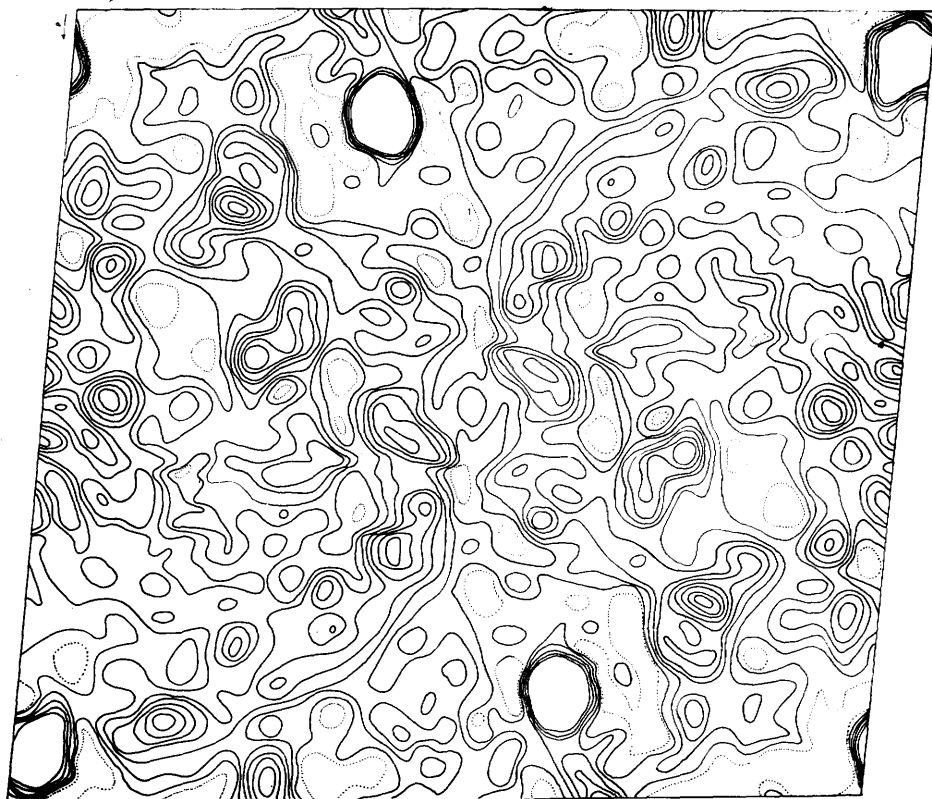


Figure 2.2. (100) projection of the Patterson vector distribution for epi-limonol iodoacetate. The iodine-iodine vectors are shown by crosses.



**Fig. 2.3a.** *The (010) Fourier projection of the electron density distribution for epi-limonol iodoacetate calculated with phase-angles derived from iodine co-ordinates alone. The zero contour is dotted. (See also Fig. 2.3b).*

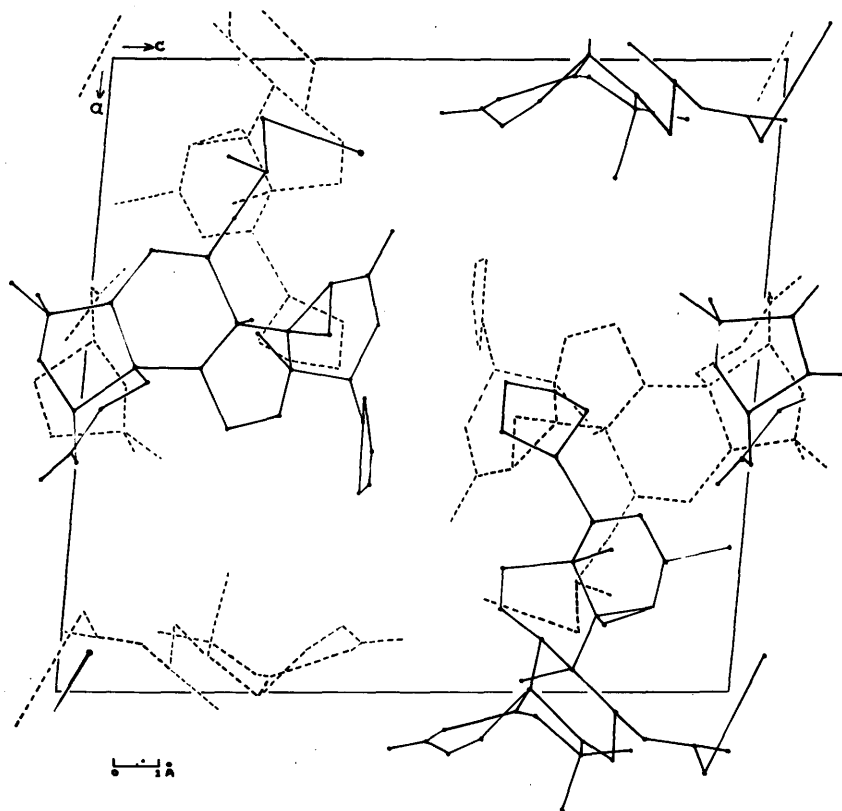


Fig. 2.3b. The unit cell contents of epi-limonol iodoacetate projected on (010).

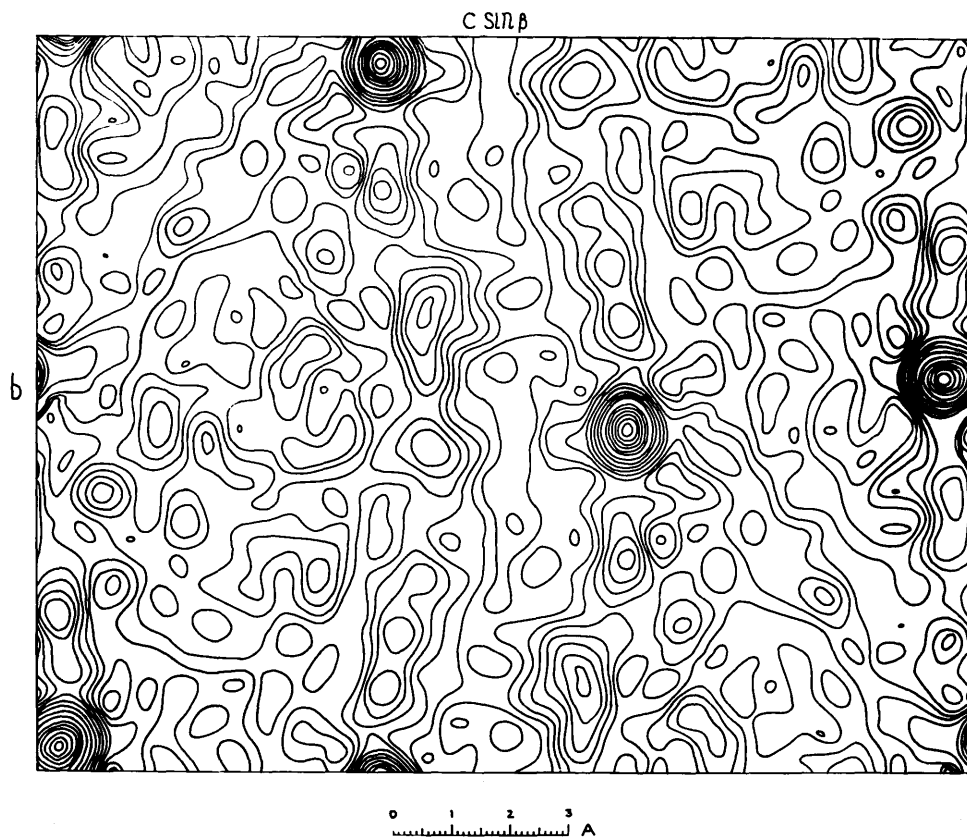


Fig. 2.4. (100). Fourier projection of the electron density distribution for epi-limonol iodoacetate calculated using phase-angles derived from the iodine co-ordinates alone.

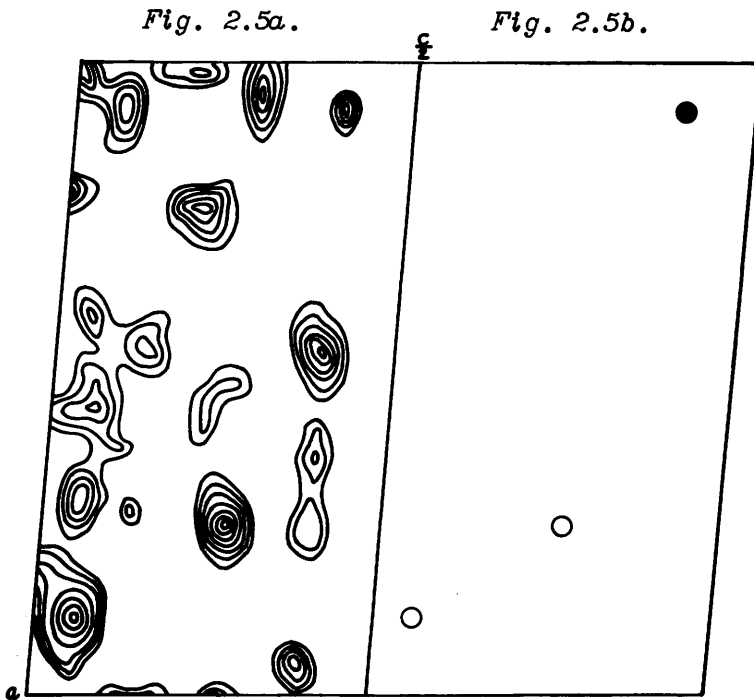
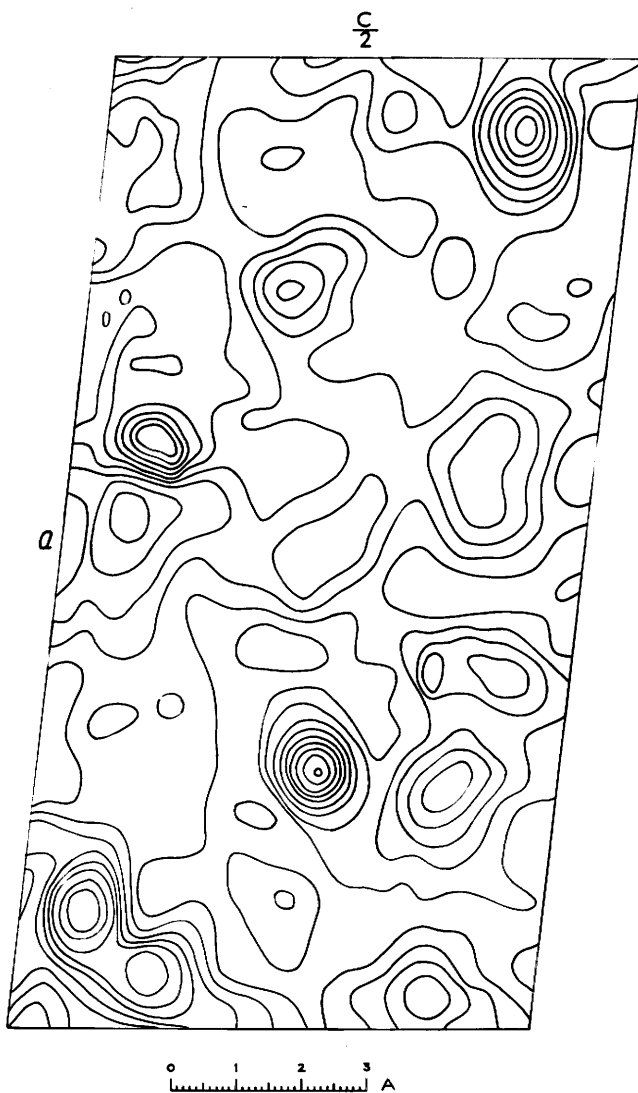


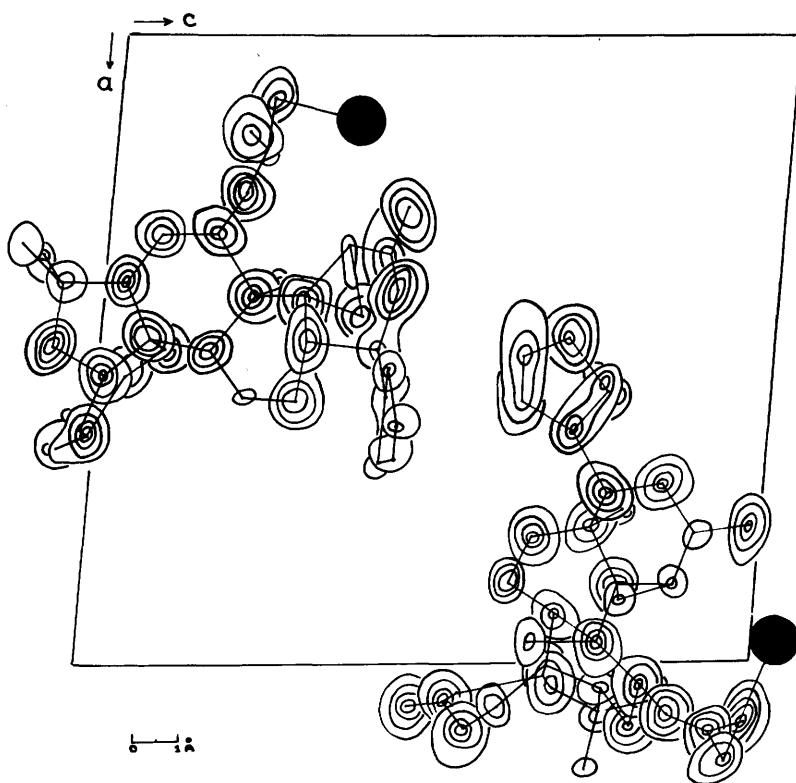
Fig. 2.5a. A Harker section at  $y = \frac{1}{2}$  for epi-limonol iodoacetate.

Fig. 2.5b. This key shows the true Harker iodine peaks (circles) and the psuedo-Harker iodine peak (shaded circle).

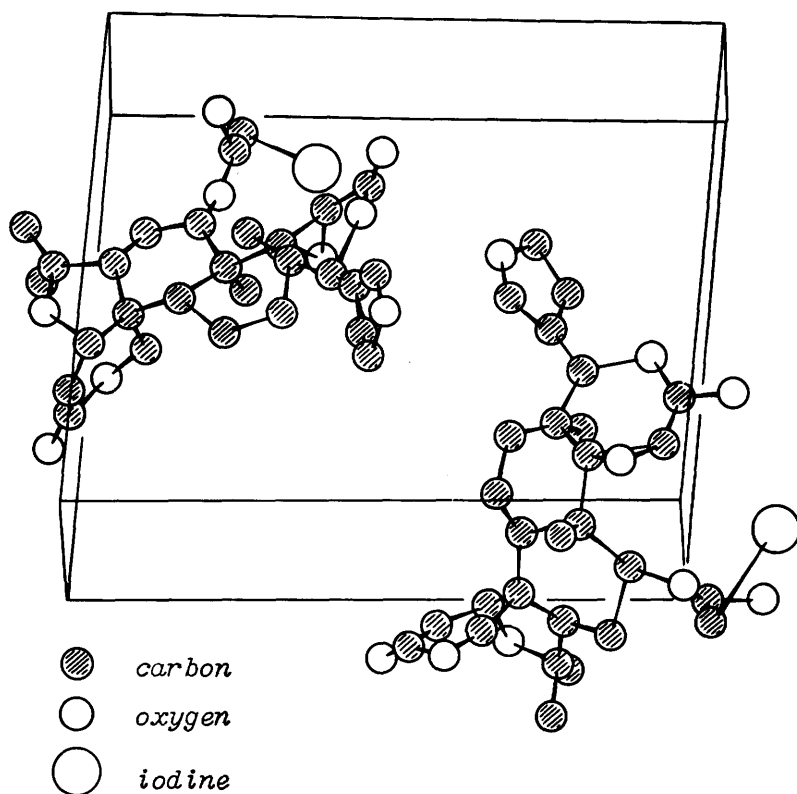


*Fig. 2.6. The Harker section at  $y = \frac{1}{2}$  for epi-limonol iodoacetate. This is a section of the final three-dimensional Patterson vector distribution calculated using all the 3,000 experimental  $|F|$ 's*





*Fig. 2.7. The two molecules of epi-limonol iodoacetate in the asymmetric crystal unit shown by means of superposed contours.*



*Figure 2.8. Model of the two molecules of epi-limonol iodoacetate in the asymmetric crystal unit showing the relative dispositions and conformation of each molecule. (The molecule depicted on the right is described by the primed co-ordinates in Table 2.3.)*

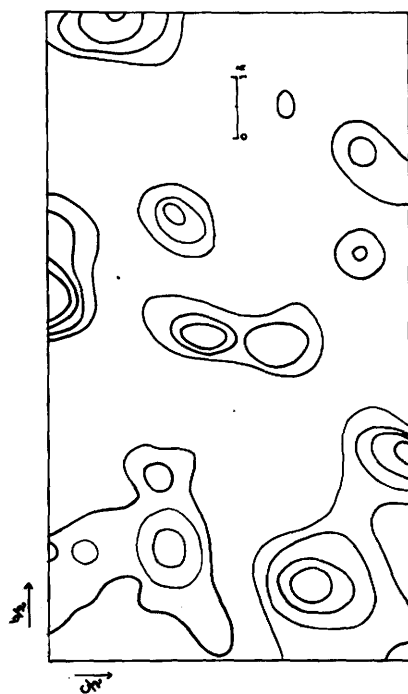


Fig. 3.1. The Harker section at  $x = \frac{1}{2}$  of the three dimensional Patterson vector distribution of neomethymycin methylene dichloride solvate.

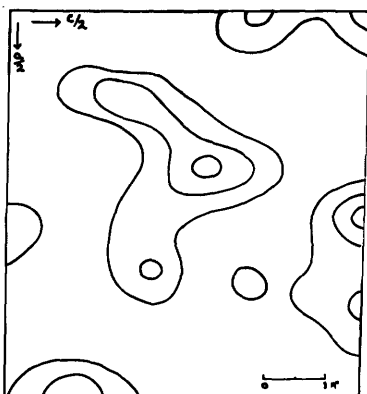


Fig. 3.2. The Harker section at  $y = \frac{1}{2}$  of the three dimensional Patterson vector distribution of neomethymycin methylene dichloride solvate.

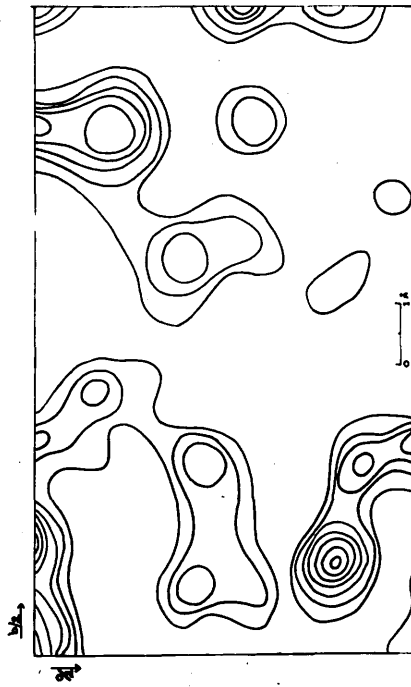
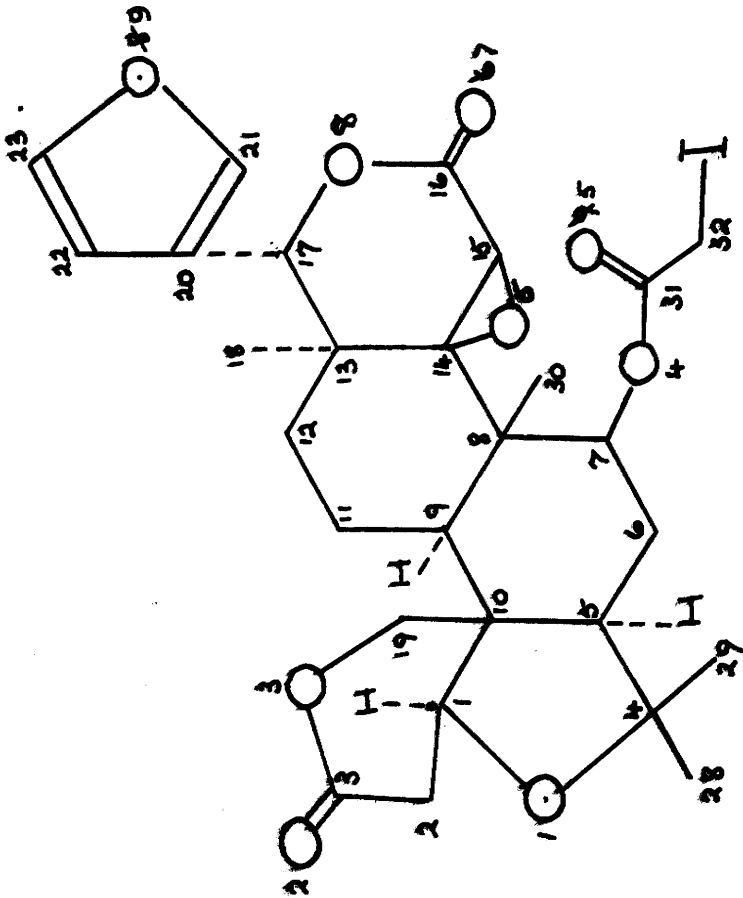
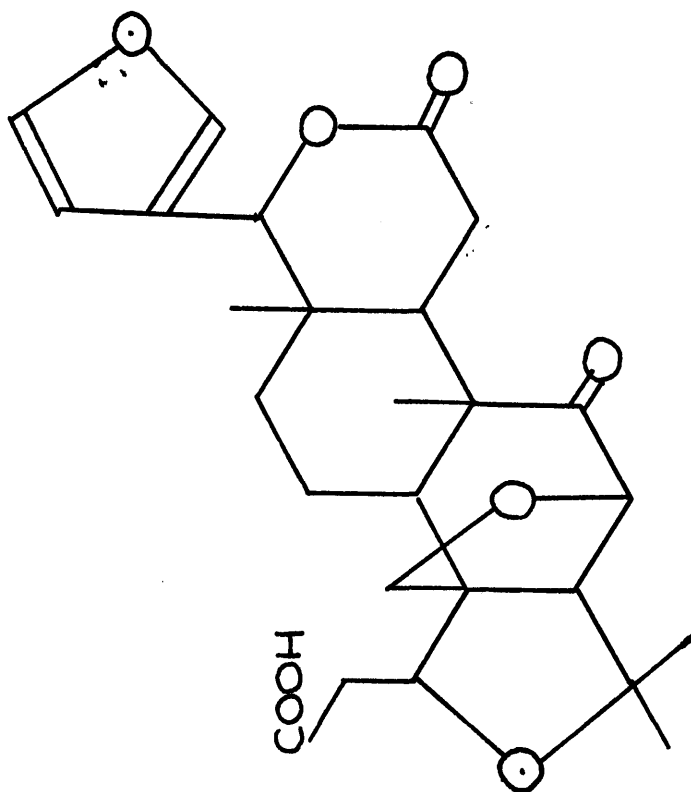


Fig. 3.3. The Harker section at  $z = \frac{1}{2}$  of the three dimensional Patterson vector distribution of neomethymycin methylene dichloride solvate.

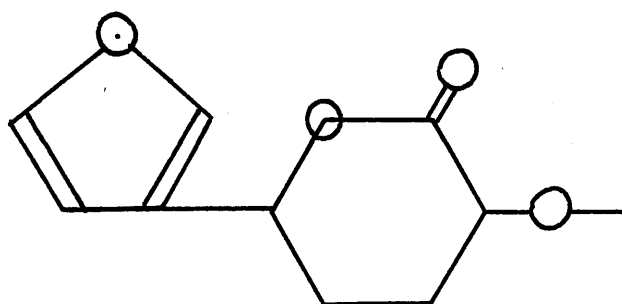
APPENDIX IV.

Structural Formulae.



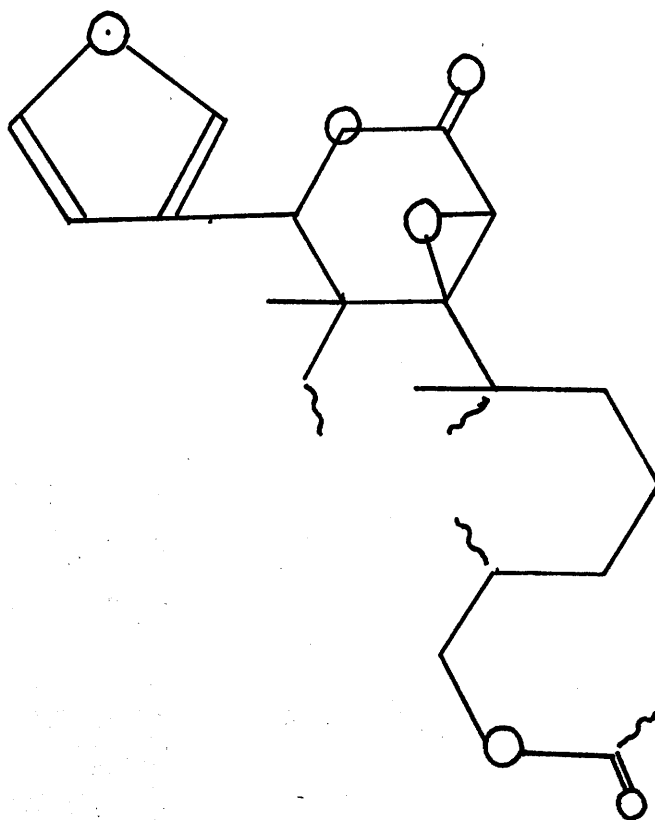


(iii)

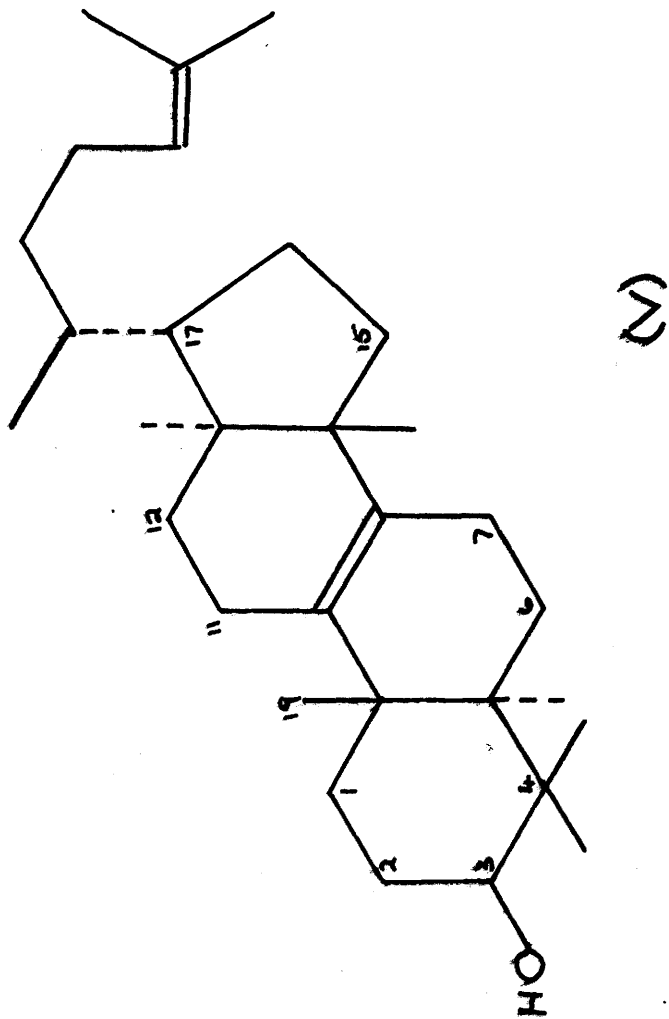


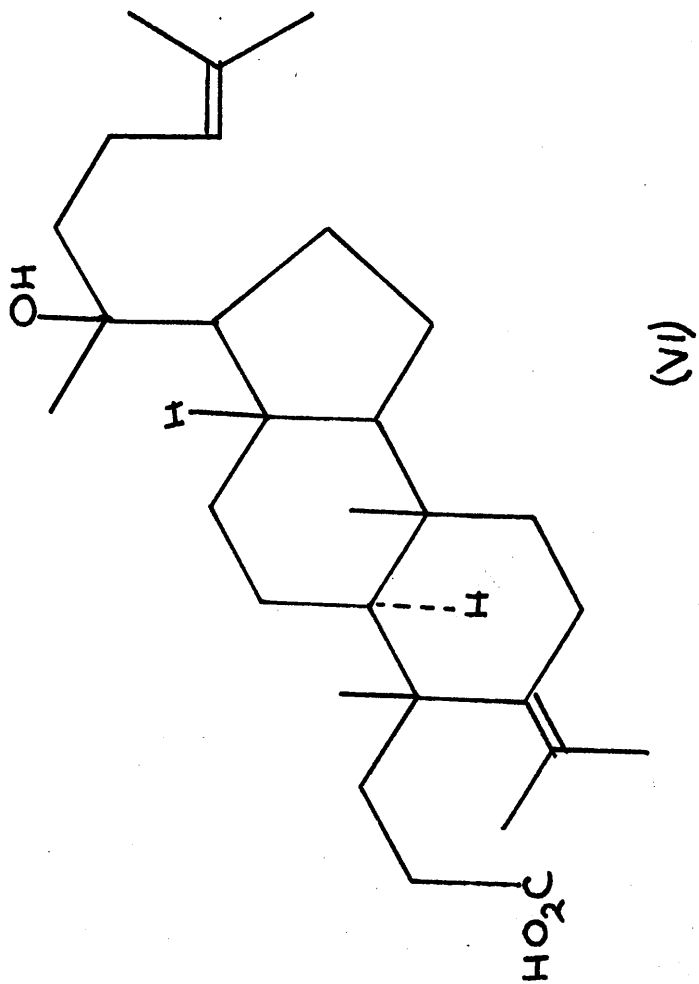
(ii)

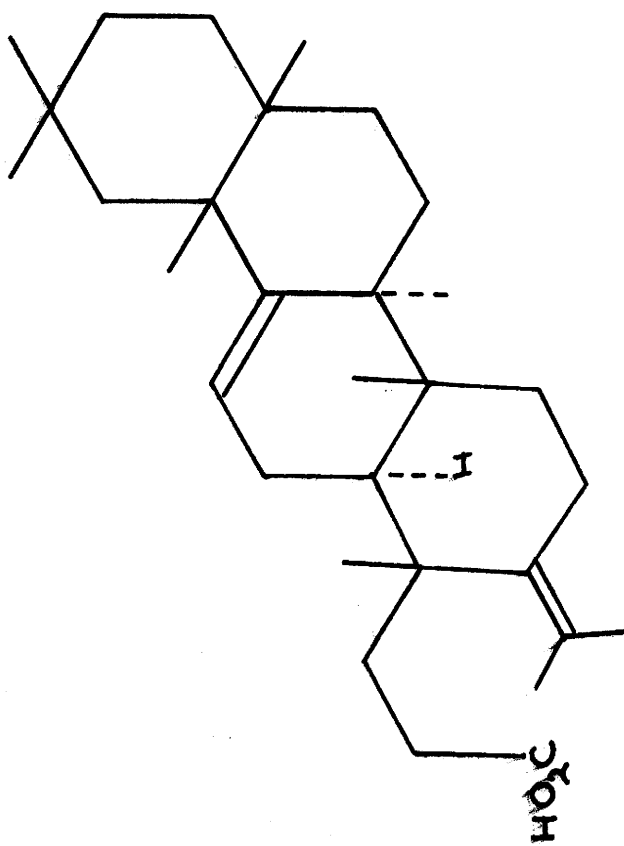


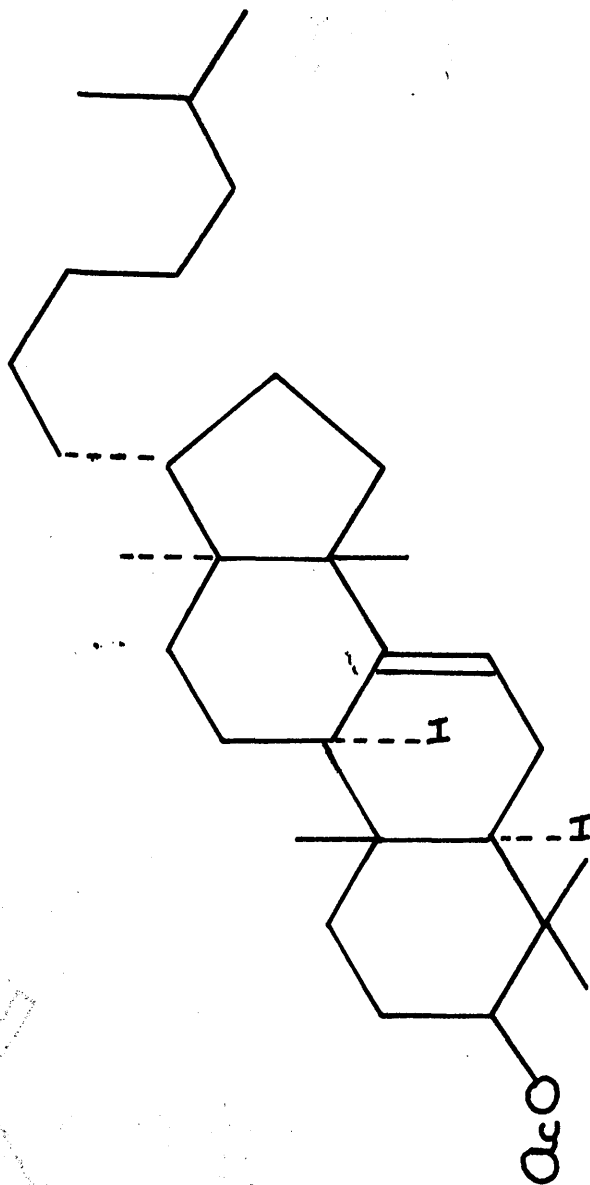


(IV)

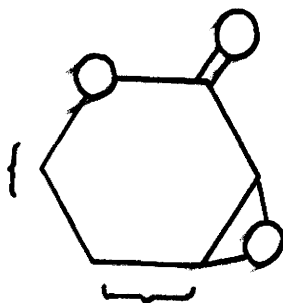




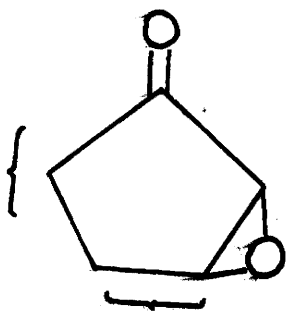




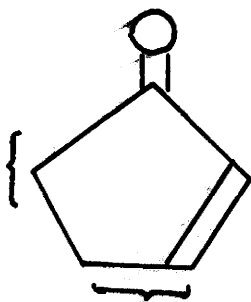
(viii)



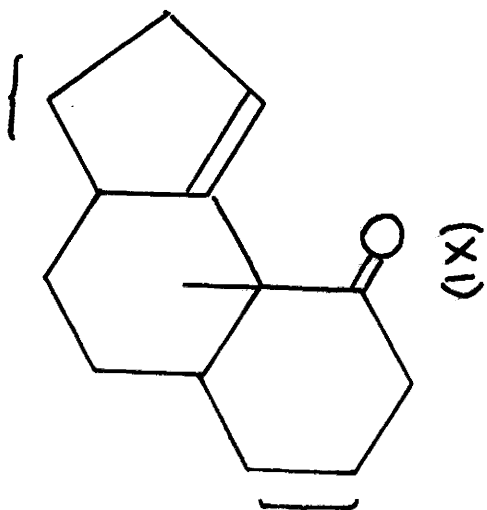
(xii)



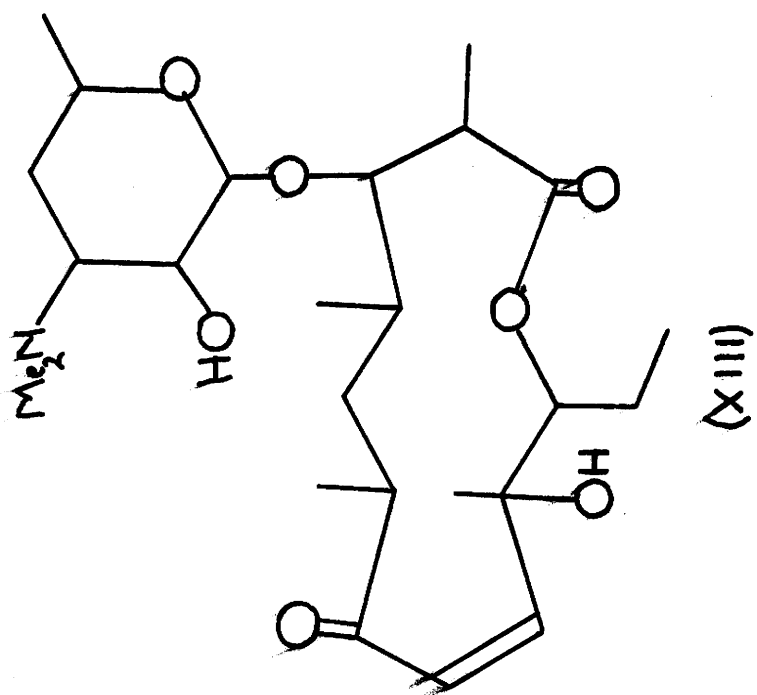
(xi)

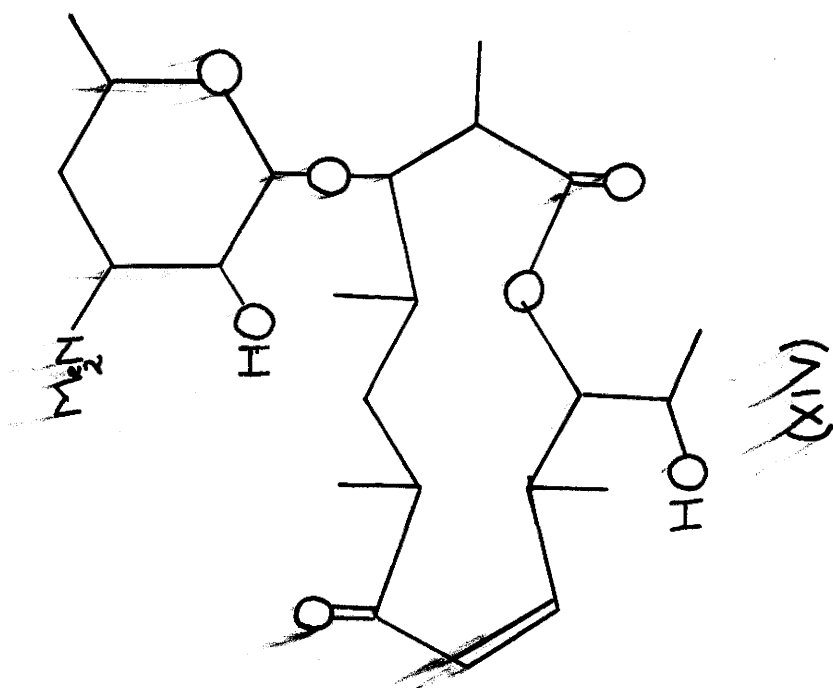


(x)

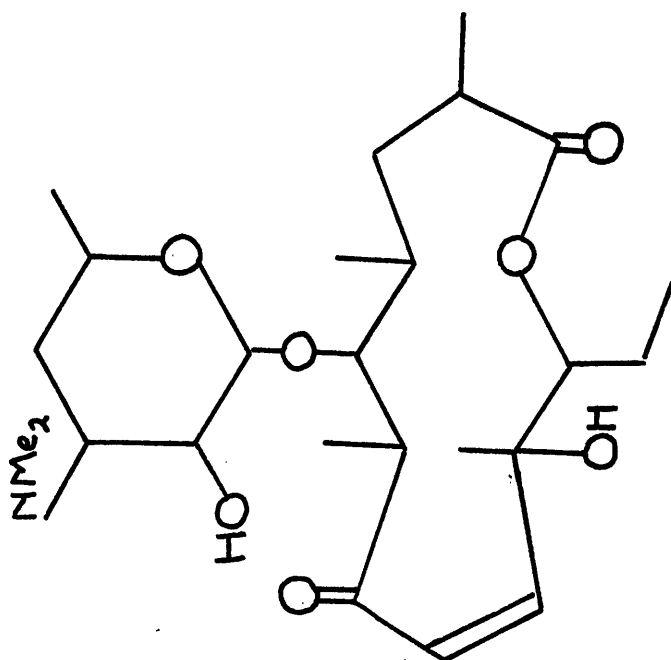


(ix)

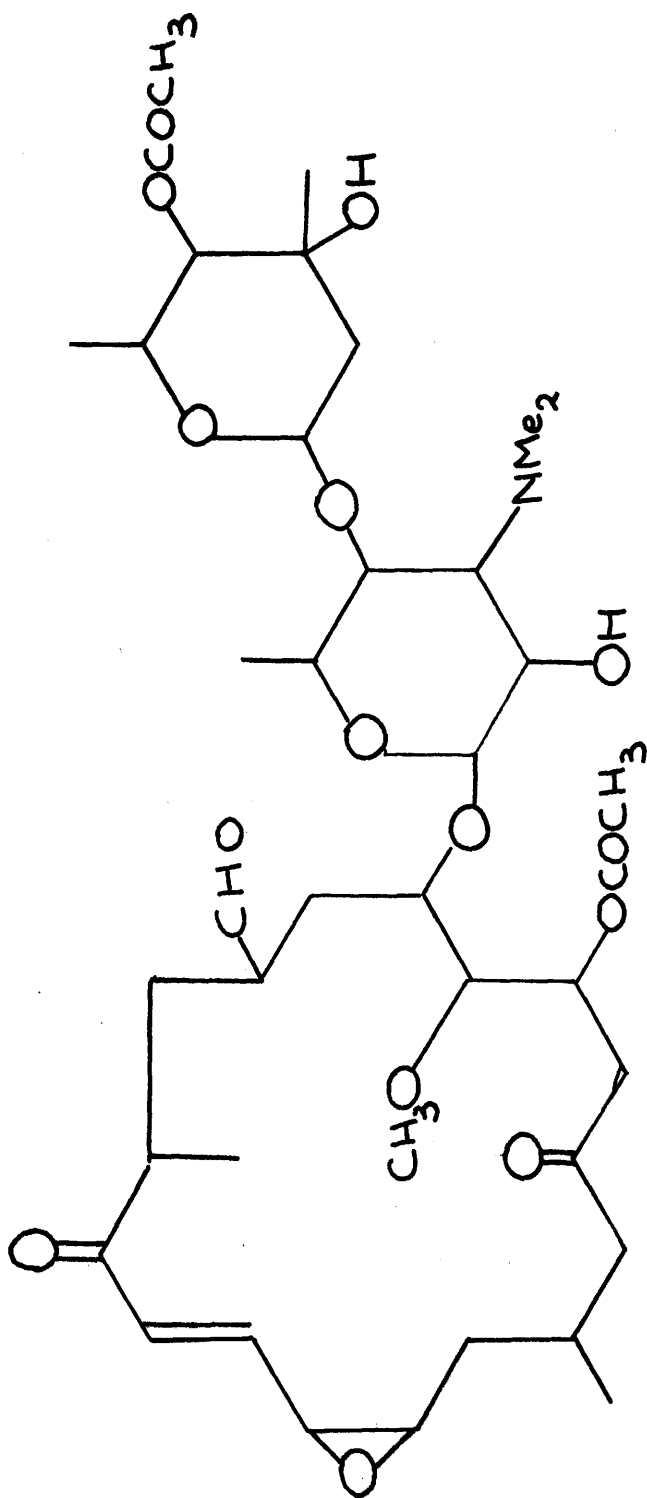




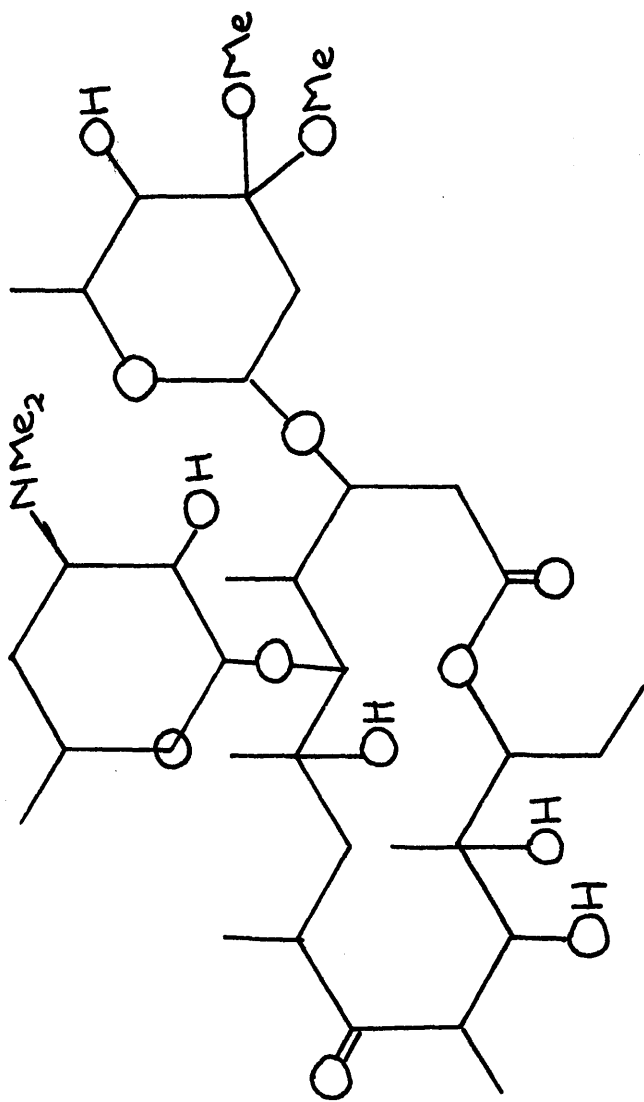




(XV)



(XVI)



(XVII)